

19 August 1982

Nuclear LUAnacy or sensible choice?

The policy of launch under nuclear attack seems to be finding increasing favour in unexpected quarters. But can we rely on machines to push the button?

Long derided as highly dangerous, the policy of launch under attack (LUA), in which one superpower would launch its land-based nuclear missiles on certification of a nuclear attack by the other, is enjoying new respectability in Washington and Moscow. This heightened interest is important because LUA has been regarded as so crazy for so long by so many and because it may well be the way in which President Reagan and Mr Caspar Weinberger, the US Secretary of Defense, end up solving the problem of how to protect the US land-based missile force. Indeed, in the future, LUA may be a more reasonable strategic option than it is now or has been — but the Department of Defense and the arms control community should study it carefully before deciding to proceed down that path.

At present the United States does not tell the Soviet Union outright whether it would launch upon warning of attack or not. And in light of the number of errors which the warning and command system could produce, and the consequent difficulty the President would have in making rapid decisions, a US policy of LUA hardly seems practical. On the other hand, the annual defence posture statements setting out US military policy, during the term of former Secretary of Defense Harold Brown, contained explicit references to the possibility that the Soviet Union must consider the likelihood of LUA before deciding to attack the United States. These hints were doubtless read with interest in Moscow.

Now, however, the issue is shifting away from official verbal hints and the obvious problems of the present system towards a discussion of future US goals. The new issue is whether, in future, the United States could build a warning and command system it could rely on to execute LUA policy, not merely preserving LUA as an option.

The interest is coming from some surprising quarters. In early July, Soviet defence minister Dmitri F. Ustinov suggested — without using those exact words — that the Soviet Union might move to an LUA policy. Ustinov was echoing what another Soviet analyst had told a group of visiting Americans. Last year, the Office of Technology Assessment of the US Congress offered a surprisingly bright view of LUA as a possible alternative to mobile basing of the MX missile. Finally, Dr Richard L. Garwin, a well known defence expert for IBM, advocates LUA. Among other LUA advocates is Harold Agnew, former director of Los Alamos National Laboratory, who is now president of the General Atomic Company. But Mr Weinberger, although continuing to mention the possibility of LUA in his posture statements, has not shown any real interest in such a policy.

Ustinov and the other Soviet official who hinted at a possible Soviet LUA policy were apparently intending to counter US nuclear weapons modernization. If you get a hair-trigger ability against us with your MX and other programmes, they seemed to say, we will be forced to respond rapidly, with LUA. Another interpretation, offered recently by Paul Bracken of the Yale School of Organization and Management, is that the Soviet Union feels it was backed in a corner by President Jimmy Carter's Presidential Decision Memorandum 59, which in 1980 shifted US policy towards fighting out a nuclear war through the macabre idea of "limited nuclear war". If the United States makes a target

of Soviet command authorities rather than Soviet cities, the Soviet Union has to consider launching its missiles early while it can still command them before the US missiles could reach the target.

There is, in fact, a debate about whether the Soviets could execute a LUA policy, no matter what they say. Soviet rockets are liquid fuelled, and cannot be kept ready to be fired at a moment's notice, unlike US rockets which are solid-fuelled. Nor is Soviet command and control on a high state of alert, normally, (partly, it is said, to prevent Soviet generals from commandeering their nuclear weapons as part of some coup). On the other hand, since any nuclear attack is likely to start during a period of general high tension, these drawbacks may not obtain, and the Soviet Union may have some chance, it is argued, of bringing off an LUA. Clearly a Soviet LUA policy, if only a declaratory one, should worry US leaders: among other things it would make Western Europe even more impressed by the Soviet bear.

The possibility of a US policy of LUA is more interesting, because a practical LUA capability is for the United States only several years and billions of dollars away (while the MX with its basing system would cost tens of billions of dollars). This was the conclusion, at least, of the OTA panel under Harry Woolf, director of the Institute for Advanced Study in Princeton. In a remarkable chapter on LUA, it said, in effect, that technology is not an obstacle to achieving some sort of meaningful LUA capability. "Doctrine and procedures — issues of national policy, not technology — more than anything else . . . determine the architecture of an LUA system," it said. The overall report, of course, did not favour any particular MX basing mode, or LUA. It merely presented several of the options to Congress.

This view contradicts the conventional view, based on past and present capabilities, that there are too many ways in which warning and communications could be disrupted or go wrong or simply be ambiguous. The President cannot, it has been said, have enough reliable information to make the cosmic decision in the 30 minutes afforded by an intercontinental ballistic missile (ICBM) attack or the 5–15 minutes in the case of a submarine-launched ballistic missile (SLBM) attack. Early warning satellites could be blinded by preemptive nuclear explosions in space, or mistake the signature of the missiles, or be so delicate as to trigger a false alert internally: and the government has revealed that the existing system triggers many such errors each year. Furthermore, SLBM's computer accessory equipment can set off false alarms: the DEW line radars have "seen" an attack when they were only looking at a flight of geese or the rising Moon.

The OTA report suggested that future systems could overcome many of these defects — although it made no claim that a system could be invulnerable. Early warning satellites could be moved out of crowded, geosynchronous orbit where a Soviet "space mine" could attack them. Instead they could be placed by themselves, further out, in unique deep orbits chosen at random, so there would be reason to suspect the innocence of any unidentified object that approached them. Since non-synchronous orbits would be used, there would have to be many warning satellites to assure that at any given time, one would be "watching" Soviet territory. If the satellites were high enough it would take a long



time — as much as 18 hours — for a ground-launched missile to reach them and knock them out.

The peculiar and varying orbits, and the use of “dormant” satellites as reserves, could make it very difficult for a Soviet anti-satellite system to knock them all out. In short, the warning system itself would contribute to security, by buying leaders more time.

The report also suggested that communications links could be maintained in a variety of ways: satellite communications could use millimetre wavelengths to avoid ionospheric disturbances due to high-altitude nuclear explosions. Rocket-launched reconstitution satellites could be on ready alert; existing communications satellites could be prepared to serve as back-ups. An LUA communications net would have an advantage over the alternative communications nets for other MX options, in that its most crucial period of use would be before the atmosphere was disrupted by nuclear attack.

In sum, the OTA report foresaw a future LUA system resilient enough to warn of a major attack — such as the lightning bolt-from-the-blue that the MX is meant to counter. The report implies that the President would not be faced with only a few minutes to think, panicked generals and blinking machines, but a relatively well ordered set of options: knowledge through early-warning of anti-satellite efforts, time to use the “hot line”, and sure information about the attack itself. The report went so far as to suggest a series of automatic LUA programmes for response, that might not need the President’s decision, in case he was incapacitated or unavailable.

Garwin takes this one step farther. If the United States had an announced, self-evident and reliable capability for LUA, the Soviet Union should be deterred from ever attempting to attack, he says. In attacking, the Soviet Union would have to launch its missiles all at once to minimize US foreknowledge. Thus, submarine-launched missiles would start hitting US bomber bases and US submarines in port 5 to 15 minutes after the launch, when the long-range ICBMs still had 15 minutes or more to go before reaching their targets.

All the President needs to be able to do, Garwin argues, is to certify that this attack is under way, based on information his machines can surely provide. He would then authorize a pre-planned launch of the US land-based ICBMs, his only surviving force. If the Soviet Union realized how automatic and surely this retaliation would come, Garwin says, it would be deterred from attacking in the first place, even against the vulnerable force of US land-based ICBMs. Thus for Garwin, LUA is the solution to the problem of Minuteman vulnerability.

Garwin’s proposed LUA policy, with its elements of automatic response and certainty of destruction of the attacking side thus has much in common with MAD — the 1960s doctrine of “mutual assured destruction” under which the United States (which then had less accurate weapons than it does now) pledged to destroy Soviet industry and society if it were ever attacked. One way to look at a possible LUA policy, therefore, is as a high-precision version of MAD.

Is all this technological hubris of the first order? What sane president, or sane society, would entrust the decision about its likely end almost entirely to machines? If the warning could not be made reliable, LUA could make a mockery of human judgement and morality.

But if technology has evolved to the point where the machinery could be made reliable we must think again. Should we decline to upgrade strategic warning, and decide not to give leaders the best information they can get? What sane president, faced with a possible nuclear attack, would want to turn the calculations over to rows of men with abacuses to make sure the “human element” was not lost in the final decision? Every day, jetliners carry innocent passengers across oceans and continents on automatic pilot systems, yet this is not morally outrageous, so long as the pilot is there to back up the machines.

A society that has already delegated so much to technology, ought, when it comes to make a final choice, to consider technology an ally, too.

An almighty crash

AEG brought about its own collapse, but there are lessons for all to learn

The collapse last week of the West German corporation AEG-Telefunken has been on the cards for more than a decade, but is none the less to be mourned on that account. The seventh largest corporation in West Germany, AEG-Telefunken is a monument to the optimism of the 1950s and 1960s, so that its decision to appeal to the courts for relief from three-fifths of its still unspecified debts marks the passing of an era. For much of the past twenty years, AEG-Telefunken had around its neck the albatross of its 40 per cent share in Kraftwerk-Union, the nuclear construction consortium — until Siemens obligingly bought it out in 1978. During the same period, but especially at the beginning of it, AEG seems accurately to have appreciated that high technology was waiting around every corner and to have determined to be strongly represented in them all. So was it just bad luck that it was not then possible to guess how high would be the cost of development in most of these new fields; or that technological conglomerates would become unfashionable; and that they would prove to be so hard to manage?

Hardly. Compared with Siemens, the obvious analogue, AEG has been run by accountants rather than by technologists. They have been skilled at raising extra capital — no less than DM1,000 million in 1979 from the West German banks, for example — but apparently less skilled at matching their technical ambitions to reality. Even so, the efforts they have made in the past three years to sell off parts of the business, and to form joint ventures with other corporations (as with Bosch over telecommunications) have been insufficient. The most obvious difficulty has been that AEG has not been technically outstanding in any of the fields in which it has taken an interest. The moral, for other corporations, should be familiar. A technically based conglomerate, however large, can survive in the long run only by technical excellence.

To many outside West Germany, it has been a surprise that one of the impediments to the reconstruction of AEG has been the resistance of the labour unions to the threat of redundancy, one of the reasons why the British General Electric Company last week pulled out of negotiations about a partnership with AEG in the manufacture of heavy electrical machinery. Despite the record of harmonious industrial relations in West Germany since the Second World War, it should be no surprise that unions are less compliant with their employers’ needs on the downturn than they were during the economic miracle. But the prospect now is that AEG will have to shed at least 40,000 of the 100,000 people it employs in West Germany and perhaps an even larger proportion of its 20,000 employees elsewhere. It is almost too much to ask that flesh and blood should stomach redundancies on such a scale. Represented as they are by law on the supervisory board of AEG, the unions have behaved predictably. Elsewhere governments have stepped in to avoid trouble of this kind (as at Chrysler and British Leyland).

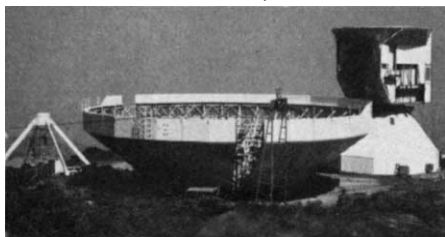
So what will be the consequences of the decline of AEG? Those who will lose their jobs, like those who have lost a large part of their investment in the company, will be more than merely inconvenienced. So too in the short run will be the host of smaller companies working as suppliers for AEG. Its competitors will now find the market for their goods more buoyant and more easily attacked. The business that AEG is forced to shed will in due course be picked up by others, with the result that some of those now destined to lose their jobs will be re-employed. In the long run, there will be two conspicuous losers. First, the legend of West German industrial cleverness will be permanently tarnished. Second, Chancellor Helmut Schmidt’s coalition government, which has steadfastly declined to intervene on AEG’s behalf (and which was politically and economically powerless to intervene) will be unfairly blamed by its opponents for not having lifted a finger and for having created the economic climate that brought AEG to its knees.

Japan reaches for the stars

Radio-dish looks set to excel

Japan is now poised to take a world lead in millimetre-wave radioastronomy — the study of molecular clouds and star-formation regions — thanks to a far-sighted decision by the Japanese Ministry of Education more than a decade ago to support the plans of Tokyo Observatory.

For while American molecular astronomers smart under the effective abandonment of their own country's plans for a large (25-metre) dish to operate in this wavelength range (see *Nature* 12 August, p.596), Professor Kenji Akabane of the Tokyo Observatory will announce this week at the International Astronomical Union meeting at Patras, Greece, his first observations with a 45-metre dish and a 5×10 -metre-dish interferometer — a combination which is likely to outclass the millimetre-wave equipment of other nations for a number of years.



Pico Veleta — a 30-m dish

This telescope complex, which took 13 years to construct and cost £20 million, is now in operation at Nobeyama, 200 km north of Tokyo.

There is only one doubt about the telescopes — that they are sited in Japan. There they are limited by a relatively humid atmosphere, which absorbs low wavelengths. Thus, said a British millimetre-wave radioastronomer this week, Japan may be restricted to measurements at wavelengths not much shorter than 1 cm, which in molecular terms means a limitation to the observations of relatively rare long molecules. Studies of isotopic abundances in the Universe — an important project for millimetre-wave astronomy — need shorter wavelengths. The Tokyo observatory, however, reckons that Nobeyama will reach down to 2 mm.

If the Nobeyama telescope fails to reach down that far there is a chance that it will be surpassed within a few years by other millimetre-wave telescopes, planned or under construction.

France, Germany, and Spain will be next, with the £15-million Institut de Radioastronomie Millimétrique (IRAM). IRAM has an administrative headquarters in Grenoble and two telescope sites: one in southern Spain, on Pico Veleta near

Granada, and one in the southern French Alps at the Plateau de Bure.

Pico Veleta is to have a 30-metre dish, whose steel-work is nearly complete — but the telescope is unlikely to be taking data before early 1984. According to Dr Dennis Downes, one of the IRAM team, this telescope should reach down to 1-mm wavelengths. But doubts have been expressed about that objective: the IRAM telescopes are to be open to the air and Sun, and so they will be distorted by wind and differential solar heating.

The exact shape of the dish must be maintained to within a wavelength of the radiation being studied — in other words 1 mm. To achieve this, the steelwork of the Pico Veleta telescope is to be heated to a uniform temperature; and the Plateau de Bure 15-metre dishes (of which there will be three making a T-shaped interferometer) will be stiffened with carbon fibre, cost

allowing. Whether these untried methods will achieve the necessary accuracy is yet to be seen. Moreover the interferometer, will take some time to construct.

Besides IRAM, Britain, together with The Netherlands, plan to build a £7-million, 15-metre dish in Hawaii. These plans require environmental approval by the Hawaii State legislature, expected towards the end of this year. The British-Dutch dish, protected by a dome transparent to the radiation, should reach down almost to the edge of the millimetre-wave atmospheric "window" at 0.35 mm. But observations are not expected there before 1986.

Thus Japan is likely to enjoy a substantial lead in this branch of astronomy, possibly the first basic science in which it might take a world lead, at least until 1984.

Robert Walgate and Alun Anderson

Remote sensing for all at UN

Vienna

Remote sensing capabilities are emerging as a keynote of Unispace-82, the Second United Nations Conference on the Peaceful Exploration and Utilization of Outer Space, now taking place in Vienna. Virtually every national delegation expressed its commitment to, and hopes for, remote sensing, and although some delegates privately expressed the views that remote sensing was still a solution in search of a problem, such sentiments were kept strictly outside of the official proceedings.

Remote sensing, for example, is to be the mainstay of the United States "global habitability" programme which is expected, during the 1980s, to develop into an umbrella for the research scheduled for the US space platform. "Global habitability" is intended to open up the entire land mass of the Earth to human habitation.

Yet the use of such techniques to survey another country contains the seeds of international tension. One of the problems to which remote sensing is repeatedly urged as a solution is the more rational use of land resources. But such techniques inevitably cut across traditional concepts of national and even personal privacy. The Soviet Union is a strong supporter of new international legislation to outlaw the improper use of data gained by satellite. As Dr Andrei P. Kapitsa, director of the Soviet stand at the exhibition associated with Unispace-82 pointed out, the publication in the Western press of predicted shortfalls in the Soviet harvest (based on satellite data) immediately raised world wheat prices, causing an extra drain on Soviet hard currency reserves.

Significantly, one of the most fervent pleas for the establishment of international principles on the use of satellite data came

from Brazil, a country frequently quoted as a classic case where only remote sensing can provide adequate survey facilities. President Joao Baptista de Oliveira Figueiredo spoke of remote sensing as an "instrument both valuable and dangerous" since it "impinges on the sovereignty of states over their natural resources".

The problem becomes even more complex with marine resources. One of the greatest advantages which Cuba derived from participation in the manned "Interkosmos" programme, according to the Cuban cosmonaut Arnaldo Tamayo Mendez, was the discovery of new fishing grounds. However, although one of the purposes of Unispace-82 was to show third world countries what space techniques are available, and although the European Space Agency has recently offered African states experimental facilities aboard Spacelab some 25 per cent of the nations represented in the United Nations have not yet reached a level of development where participation of any kind in a remote sensing survey is an economic possibility. The use of satellites to determine fishshoals in international waters could mean that those shoals fail to reach the traditional fishing grounds of the disadvantaged nations.

To obviate such inequities, some speakers at the forum of non-governmental organizations (NGOs), which took place in parallel to the main sessions, urged that the third world should strive for its own launch capability, but without clearly suggesting how this might be achieved. More practical would be the creation of an international pool of land-resource data as part of the United Nations commitment to space. The two space superpowers already have a data-pooling

programme, which is apparently continuing, although, as the administrator of the National Aeronautics and Space Administration, James M. Beggs, made it clear there can be no resumption of joint US-USSR experiments or missions until the international political climate improves considerably.

Several countries including India, France and the host country, Austria, have already established their own cooperation programmes with both the capitalist and the socialist worlds. Yet it was occasionally

difficult to pin down the delegates from these countries on the advantages their countries have derived from such cooperation. And the Polish cosmonaut, Mirosław Hermaszewski, expertly fielded the question of what benefits Poland had derived from participating in the Comecon space programme — although, only a few metres away, the Polish Institute of Cartography had on display an excellent series of maps, which synthesized data from both the Interkosmos and the Landsat programmes.

Vera Rich

South African funds

The UK Science and Engineering Research Council (SERC), is reviewing its future commitment to the South African Observatory (SAO), because of financial pressures. South African astronomers currently take up 65 per cent of the viewing time of SAO, and SERC "buys" the rest for UK astronomers; has been suggested that this proportion be decreased. A recent questionnaire, however, revealed a strong desire within the UK astronomical community for this level of involvement to continue.

In terms of technical sophistication and size, the telescopes of the SAO (near Capetown) are outranked by others available to British astronomers in the Southern Hemisphere, particularly the Anglo-Australian and UK Schmidt telescopes in Australia. Paradoxically, SAO's usefulness is increased because its telescopes are not so advanced as to be oversubscribed in demand and yet are sufficiently powerful to do valuable work. In particular, SAO is heavily used for optical and infrared photometry with such objects as the brighter active galaxies and also cataclysmic variable and Cepheid-variable stars.

Earlier this year, while reviewing its plans for the financial year 1983-84 and the following three years, the Astronomy, Space and Radio Board of SERC set up a working group to ascertain the need for continuing use of SAO. The group circulated a questionnaire, to which it received replies from 60 or so groups in the United Kingdom. SERC says the working group reported to the board in July that there was a strong user demand for SAO.

The position of the board is that the decision will depend on scientific and financial considerations and that political issues are not relevant. According to SERC, however, a handful of the questionnaires raised political reasons for discontinuing the use of SAO. One astronomer, Dr Michael Rowan-Robinson, who has just completed a period of service on the SERC committee responsible for SAO, feels particularly strongly: "I seem to be in a minority in feeling that SERC should not maintain an official link of this kind with the South African government. Use of the telescope should be left more to the consciences of individual astronomers."

SERC is due soon to start negotiating its next contract with SAO, to which it now pays about £240,000 a year (about 25 per cent of the annual budget of the observatory) for 35 per cent of the total observing time. The British astronomical community will have a chance to review its use of all observatories in the Southern Hemisphere at the Anglo-Australian Telescope symposium to be held next month.

Philip Campbell

Remote sensing satellites

French plans for spot satellites

The French national space agency, the Centre National d'Etudes Spatiales (CNES), believes there is money to be made by remote sensing from space. Sales of data recorded by the two Spot satellites, France's first national remote sensing satellites due for launch in 1984 and 1986, are expected to offset the capital cost of the satellites' development and running costs in about ten years. Responsibility for earning the revenue will rest with Spot-Image, the independent company created early last month to market Spot data worldwide.

CNES's approach to the commercialization of remote sensing data differs markedly from that in the United States. The US National Aeronautics and Space Administration (NASA), like CNES chiefly a research and development agency, has been an unwilling operator of the service forced upon it when data from the early Landsat satellites proved commercially attractive. Potential users have complained that NASA was sometimes insensitive to their needs.

NASA will indeed be administering the distribution of data from Landsat-D, the latest satellite in the series launched last month, and from Landsat-D', its successor, but beyond 1987 the Reagan Administration has insisted that NASA will have to get out of commerce so that the operational service will have to be taken up by private industry. Thus the fortunes of Spot-Image will be keenly watched in the United States as well as by the European Space Agency, now debating how to make available data from its planned remote sensing satellite ERS1, and the British government, which is contemplating a major effort in remote sensing.

Spot-Image's main shareholders are CNES with a 34 per cent holding, three French government agencies each with 10 per cent — the Institut Cartographique Nationale, the Institut Français du Pétrole and the Bureau de Recherche Géologique des Minières — and two companies, Matra and Société Européenne de Propulsion, with 7.5 per cent each. Organizations from Sweden and Belgium which have collaborated to some extent on Spot, will

also have small shareholdings. Spot-Image will buy data from CNES and sell them.

NASA's early experience in selling remote sensing data, however, seems not to augur well for Spot-Image. Although there has been great interest in Landsat data, there has been no prospect that fees would match development costs. But, according to Gérard Brachet, director-general of Spot-Image, the market for data from the second generation of remote sensing satellites should grow by leaps and bounds — he reckons that Spot should earn about \$900 million, to cover development and running costs, within ten years.

The chief advantages of Spot and Landsat-D over earlier generations of remote sensing satellites is their increased resolution and number of spectral bands. Spot, for example, whose chief instrument is a multilinear array, will observe in three bands in the infrared and visible regions of the spectrum with a ground resolution of 20 m and in a broad band with 10 m resolution. Stereoscopic images are also available.

Spot data will be received and interpreted in Europe at ground stations in Toulouse (France) and Kiruna (Sweden). Spot-Image plans to turn round data within 48 hours for dispatch to users. The charge for one scene in digital form is expected to be about \$1,000, but the fee for a photographic image has yet to be worked out, although it will be less. Brachet is confident that prices will compare favourably with those for images generated by the multi-spectra scanner aboard Landsat-D although comparisons with the thematic mapper, Landsat-D's most innovative instrument, are more uncertain.

Spot-Image is negotiating with some countries to receive Spot data direct via their own Earth stations, in which case there will be a copyright fee of only about FF220 (\$40) per scene, less than the cost of images bought from Toulouse or Kiruna. The apparent discrepancy is explained, according to Brachet, by the prospective loss of about 70 per cent of potential images due to cloud as well as the cost of processing images.

Judy Redfearn

French Comité National

Democracy, confusion abounds

One of the major and unique institutions of French scientific democracy, the Comité National, a kind of scientific parliament with effective power over the work and careers of thousands of French researchers, is to be transformed.

That fact alone would lead to a great deal of heat, given the generally polemical French nature. But the fire has been stoked even higher by the decree, just published, which describes the transformation. It leaves so many loose ends that many French scientists remain unclear whether they can vote for candidates to the new parliament or not.

This is important, because of the potential power invested in the Comité National: if a scientist can vote for a candidate, at least he or she is assured of some kind of representation.

The Comité is the 1,200-strong board of assessment of the Centre National de la Recherche Scientifique (CNRS). CNRS is only one of the principal French research organizations, but it is the biggest and arguably the most important — at least for basic science; this year CNRS controls a budget of more than FF 6,000 million (over £500 million), and supports 9,322 scientists and 14,514 engineers, technicians and administrators. In general, the best of the university laboratories are at least "associated" with (or partly supported by) CNRS. And within this organization the Comité National plays a role by giving its advice (which is usually accepted) to the CNRS research directors on such matters as the hiring and firing of staff, opening and closing laboratories, and the awarding of grants.

The Comité National may appear to be no more than a collection of peer review committees — and it certainly does function in that way, divided into 45 sections each of 25 people according to subject or research. But the Comité is unique because it is an *elected* body, with a French electorate now totalling perhaps 20,000 people; and the key to the squabbles over the Comité is the question of what groups will be represented and to what extent?

Jean-Pierre Chevènement, minister for science and industry, was unhappy with the Comité that he inherited (the last election to the Comité National was in 1980; President Mitterrand came to power in mid-1981) for two main reasons — its subject structure was out of date, and it lacked representation from the technicians and administrators.

The new decree for the Comité sorts out these matters, and others besides. Chevènement has slightly increased the number of Comité seats that he can name himself, on advice from the CNRS directorate. He can now name eight rather

than nine of the 25 members of each Comité (the rest being elected); and four seats per section are now reserved for election by engineers, technicians and administrators. The remaining 12 seats per section (just under half) are to be determined by election by scientists. But how is a scientist to be defined in that context? It is here that the new decree is unclear, and it is here that there will be plenty of discussion and acrimony before the next elections to the Comité National, planned for early 1983.

For the new decree actually decreases the right of certain university researchers to an automatic vote. Previously any researcher could vote; now only researchers with a "link" (the word is deliberately ambiguous) to CNRS may vote. This may be quite reasonable — after all, the Comité directly affects only CNRS employees — but CNRS is so important that researchers were pleased to have their little right to "meddle" in CNRS affairs.

Now that right seems to have gone. Or has it? According to the decree, scientific institutions that are not part of the CNRS may still appoint certain of their staff to vote for the Comité. Among those institutions could be universities. So there may yet be a back door to a voting right. And certain other categories of people do not have an automatic vote, but may receive one if they themselves apply to CNRS for the right.

What this means in effect is that almost every scientist and technician in France will have some way of getting a vote for the Comité; but for some a vote will come more easily than for others. It seems the ministry hopes that this solution will reduce political argument about rights to vote, but at the same time it will be a hard winter for the CNRS elector committee which, between now and January, will have to decide exactly who can vote and why. This process alone will take fully six months, CNRS estimates.

Robert Walgate

US computer industry

Paying the price

Washington

Federal Judge Harold H. Greene has raised the price that American Telephone and Telegraph Company (AT&T) will have to pay for the privilege of entering the data processing and computer game.

Ruling on 11 August on the proposed anti-trust settlement between AT&T and the Justice Department, Judge Greene said he would accept the basic deal, under which AT&T gives up its local telephone companies in exchange for the right to enter the unregulated computer market.

But he insisted on certain modifications. The local telephone companies should be allowed to keep publishing the Yellow Pages directories, which are a big money-maker; they should also be permitted to market, but not manufacture, telephone equipment, he said. Under the original settlement, both of these options would be reserved for the parent AT&T company.

The judge also insisted that AT&T be barred from entering the "electronic publishing" field for at least seven years. This excludes AT&T for the time being from a variety of electronic information and news services; newspaper publishers have been especially worried that AT&T's grip on the country's communication system would give it an unfair advantage in this fledgling industry. Under Judge Greene's proposal, AT&T apparently could still supply transmission lines and terminal equipment for such ventures, but could not do the actual collection and compilation of information.

The anti-trust law limits the judge to making suggestions; he cannot order changes in the settlement. He can, however, reject it, and Judge Greene did not mince words: if the parties do not agree to his "suggestions", he will throw out the settlement and reopen the anti-trust case — which has already dragged on for eight years.

AT&T's vice-president and general counsel, William Keefauver, said "AT&T has a strong incentive to accept a decree and free ourselves from the business restrictions of the 1956 decree". (The 1956 settlement barred AT&T from entering the unregulated computer market. It resulted from earlier charges that AT&T was using revenues from its monopoly telephone business to subsidize its competitive ventures.) Failure to accept Judge Greene's terms means that those barriers remain. Keefauver said that the judge's suggested modifications "don't dramatically impact the thrust of the decree".

The Justice Department is less certain to go along with the changes. It had demanded that local companies should not market telephone equipment — a competitive business — while operating as regulated monopolies. The judge ruled that this was merely "theoretical consistency", when in fact allowing the companies to market equipment would increase competition — and at the same time keep rates down.

AT&T and the Justice Department have 15 days to respond to the judge. Earlier this year, Representative Timothy Wirth (Democrat, Colorado) introduced legislation to stiffen the terms of the anti-trust settlement; he later withdrew it in the face of heavy lobbying by AT&T.

The judge's suggested changes appear to incorporate a substantial portion of the Wirth plan, in particular letting the local companies keep the Yellow Pages and the right to market equipment. But, significantly, the judge did not recommend

the other major point in Wirth's package: breaking off Long Lines (AT&T's long-distance network, which holds a virtual monopoly on interstate communications) from the more competitive divisions of the company. AT&T would thus retain Long Lines, Western Electric (which manufactures communications equipment), American Bell (the new subsidiary entering the computer market), and Bell Labs without restrictions on how research and development funds are allocated.

Stephen Budiansky

US biotechnology

Re-entry plans

Washington

E.F. Hutton, the large US financial company, is making its second entry into the potentially lucrative field of founding and backing fledgling biotechnology companies. In September, Hutton will begin offering investors limited partnerships in California Biotechnology Inc., a new company organized around the talents of three prominent university researchers. Hutton has already grossed \$5.4 million in startup capital for the firm, which is now building a laboratory in Mountain View, in the San Francisco area.

Cal Biotech, as the firm is called, will have as director of research Professor Brian J. McCarthy on leave of absence from his position at the University of California, at Irvine. John Baxter of the University of California at San Francisco, has agreed to consult exclusively for Cal Biotech, although he will retain his professorship at the university. The third star is to be John Shine of Australia National University at Canberra. Shine will also consult exclusively for Cal Biotech, and work for the company under contract at his lab in Canberra, according to Hutton official Zsolt Harsanyi.

The company plans to use DNA techniques to develop pharmaceutical drugs, which could be tested and marketed by major pharmaceutical firms. Cal Biotech's role will be limited to development and ownership of the drugs themselves. Initially, the group plans to pursue development of human pharmaceuticals including those useful for cardiovascular diseases and anti-inflammatory purposes. However, Baxter and Shine have developed a way to produce beta-endorphin, the patent for which is held by the University of California. Baxter's laboratory has also cloned human and bovine growth hormone gene, so the company's work could proceed in those areas as well.

"What makes our company different from other biotechnology companies" says the firm's president Alfred G. Scheid, a long term consultant and former Hutton employee who makes rather a speciality of founding new companies, "is the

combination of top scientists and access to a continuing flow of capital." Besides sponsoring the company financially, Hutton has also lent one of its executive vice presidents, William G. Baker, to be chairman of the board. Hutton itself invested an undisclosed amount as part of the \$5.4 million fundraising.

Harsanyi says that come September, investors can put up a minimum of \$5,000 to become limited partners in Cal Biotech. Using a little-noted provision in the tax law (that has been available for research and developments partnerships since the mid-1970s), they will be able to deduct perhaps as much as 99 per cent of the money they invest if the company spends the money in that year. Thus, the attractiveness of the investment depends on Cal Biotech's financial planning, which may explain why Hutton is taking such a deep interest in its management. Besides the tax writeoff, investors will receive a share of any instant royalties and profits the company earns from drugs that are sold, which could start as early as 1988, Harsanyi says. The company itself will be the general partner, splitting profits and royalties with the group of limited partners.

The approach of forming a company around a group of researchers and limiting its scope to their research is a major switch for E.F. Hutton following its failure with an alternative approach. In February 1981 Hutton raised \$40 million to start DNA Science, a company planned to sponsor biotechnology research in many institutions around the world. Most of the initial investment was to go to the Weizmann Institute in Rehovot, Israel, to support work directed by Christian B. Anfinsen who had moved there after retiring from his post at the US National Institutes of Health. Baxter's laboratory in California was also due for support.

Research funding by DNA Science was mixed up with possible marketing rights granted to two major firms, and the whole caboodle was to be managed by a businessman, E. Russell Eggers, with two Hutton officials, Harsanyi and Nelson Schneider, as vice presidents. But the structure of the company was too unwieldy and the 1981 tax law cancelled out some of the expected tax benefits from investment. So on 4 August 1981, the investors got their money back.

The lesson of DNA Science, Harsanyi says, is that investors in biotechnology are attracted by key people, such as Anfinsen and Baxter. The most promising approach, therefore, is to structure a company around these people, not expecting them to do management and marketing but assuring their access to capital for research and development. Whether Cal Biotech can develop a group of important products, to see them through the hurdles of trials and testing, and bring its investors golden returns, remains to be seen.

Deborah Shapley

Biotechnology centre

Links between industry and biotechnology research are thriving at the University of Leicester. Four companies are putting up approximately £1 million between them to establish a new biotechnology centre at the university. And the Science and Engineering Research Council (SERC) has promised the new centre £180,000 for capital equipment.

Industry's interest in Leicester is particularly timely. The University Grants Committee (UGC) recently awarded the university £50,000 of its grant earmarked for biotechnology to create three new lectureships. The university is now hoping to persuade the grants committee that one of those posts, for a yeast specialist, should be created within the new biocentre. The others are reserved for a plant biotechnologist and a mammalian geneticist within the university proper.

The companies prepared to put their money into research at Leicester are Whitbread, the brewers, Gallaher, the tobacco company, Dalgety-Spillers, the food manufacturers and John Brown Engineers and Constructors whose main claim to biotechnical fame is the construction of the Pruteen plant for ICI. They have guaranteed support to the new centre for five years and have already agreed a programme of research. The centre will be concerned primarily with recombinant DNA technology, although questions of scale-up may be considered later. The research programme will consider plasmid DNA regulation and protein secretion in yeasts and the analysis and structure of genes in higher plants.

Money from the four companies should be sufficient to keep about eight researchers employed for five years, but the centre will also be trying to build up a sound contract research business. Patents resulting from work carried out under the core programme will be shared between the four companies and the university. The share of revenue will depend on how the patent is licensed. Ultimately, the centre hopes to build up a contract research business, the profits from which will be ploughed back into the centre. Precisely how much the university will earn from its association with the centre remains uncertain.

Even if the university has little to gain financially, it is intended to benefit from the small teaching commitment that the centre will take on. Training is also to be provided for employees seconded from industry. Initially, the centre will be housed in a suite of laboratories in the university's pre-clinical medical sciences building (for which it may pay no rent), but it may later build its own accommodation if cash can be raised.

Judy Redfearn

New sights for Land

Washington

Dr Edwin H. Land, a legendary figure in US industrial history for his invention of the instant camera and for founding the Polaroid Corporation, announced on 27 July that he would resign as a director and chairman of the board of the company. When a remaining role as a consultant expires at the end of the year, Land will have relinquished all direct responsibility for the company identified with his name. Since 1975, he has been gradually breaking his links with the corporation in order to devote himself to research on colour vision and related fields. He is now 73.

Freedom ideal

As an inventor, however, Land has surprised no one in inventing an entirely new setting for his new career as a full-time scientist. Earlier this month, he gave me an exclusive tour of the Rowland Institute, a luxurious new building of brick, slate and oak that he has built and endowed on the banks of the Charles River in Cambridge, Massachusetts, only a few blocks from his old centre of operations at Polaroid.

At present, only 25 of the 80 people expected to work at the institute have been engaged, but Land is in no hurry to fill it up quickly, he explained, as he mounted, two steps at a time, a circular stairway carved in poplar wood leading to a large two-storey atrium housing a Japanese style garden and off which lay two storeys of eight double-laboratories.

Land's favourite word to describe the idea of the institute is freedom. Here, he and colleagues will be free from the pressures of writing government funding proposals or the demands of industry — or, perhaps, the burden of running Polaroid, which has had its troubles in recent years. To ensure this freedom, Land has endowed the institute with \$20 million. "This is an experiment to see what will happen if a group of competent, intelligent and bright people were pursuing the substance of science itself for its own sake, without thought of product or profit", he says.

Some of the basic work to be undertaken, such as the study of light-induced heat transfer across molecular surfaces, is not far from work done at Polaroid, which was famous for exploring the frontiers of the science that underlay its projects. "We were very proud of the basic work we did at Polaroid. But while we were doing it we carried a very different and very proud burden.

"It was a joy to make new films, new dyes and new cameras" said the holder of 533 patents, who is a member of the US Inventors Hall of Fame, along with Thomas Alva Edison and others. "It is a different kind of joy if the product is deep insight into natural phenomena from which there may be by-products internally". The institute, like any other

non-profit group, will be eligible to apply for patents and to accept royalties. But, Land stresses, basic science will be the product, with any inventions that transpire the by-product — not the other way around. Perhaps the Rowland Institute is Polaroid in reverse.

The institute aims to be collegial in style, explains James Foley, an organic chemist from Polaroid whom Land hired nine months ago. "Each of us is a sun on our own project, with the others as satellites, and I will be a satellite on Jim's project" says Land. Projects at present contemplated range from Land's own work on the physics of colour vision to an examination of the structure of biologically important molecules.



Edwin Land, new perspectives

For all its newness, its large, quiet library, Mexican grey pebbles and whisper-perfect auditorium, the institute is avoiding stocking up on fancy equipment (although each scientist may have a computer terminal in his office). Nor is Land hiring "teams" of people to work on a single project. He is opposed to both as too likely to drag researchers away from pursuing a clear, logical line of enquiry. Once you have a big machine, he says, you tend to work on problems the machine can do, rather than those that are most important or interesting.

The enormous darkroom in the basement, for example, has only a few items in it. There, Land happily shows off a phenomenon he discovered only the day before, in which the eye, after exposure to a certain kind of light, sees the fan-shaped spokes on a wheel become narrower. "We're going to look at one thing, then the next, and the next, without trying to explain too much at first. In a sense, we're returning to an older style of doing science. This was how Faraday worked."

Base of success

Back at Polaroid, the reins of power are now firmly in the hands of William J. McCune Jr who has been with the company since 1939, and whom Land made president in 1975, chief executive officer in 1980, and who now becomes chairman of

the board too. McCune, who is 67, is an independent Polaroid power in his own right. One story has it that he recently rebuffed two would-be kidnappers in a local parking lot. Most discussions of Polaroid's health and future tend to show McCune and the board of directors seeing things differently from Land, although there is no indication of an outright break being the cause of Land's present departure.

Indeed, Polaroid, while highly successful with sales of \$1,400 million, has had several things to disagree about internally. The film manufacturers have been hard hit by the rising price of silver in the late 1970s. In addition, the amateur camera market, which accounts for 80 per cent of Polaroid's business, levelled in those years in response to the recession, at a time when Polaroid had doubled its camera and film-making capacity in what proved to be a mistaken expectation of continued growth. Both US sales and sales in Europe, which had grown spectacularly, tailed off. The company workforce, which numbered 20,000 in 1979, now stands at approximately 15,000, and a number of high executives have been offered generous terms to leave to reduce management too.

Finally, the company has gone on introducing the kind of spectacular, unique inventions for which Land was famous — such as Time Zero film, an improvement on the SX-70 film first introduced in 1972. But it has also recently experienced its first real failure: Polavision, an instant movie camera and film system introduced in 1977 and withdrawn in 1979. The system came just when video products were coming along, and cost more than a comparable video system; the company had to write off an estimated \$68 million and opinions vary over the extent to which Land bears the blame. McCune is seen as strongly favouring the company's decision, in recent years, to expand out of the volatile amateur camera market and into commercial, industrial and scientific products. Polaroid now offers cameras for computer graphics to industry and cameras for medical purposes, including nuclear medicine for example.

Will Polaroid go on being inventive after Land? He says "One of my functions was — innocently — to insulate the company from many of the conventional criteria by which companies are judged on a monthly or quarterly basis.

"If with my insulation removed, the company becomes susceptible to these judgements . . . this will lead to a basic conflict between two different approaches: first, the long-term, imaginative approach of sensing an unmet human need and bringing it, by way of science, to fulfillment; and second, responding to a chorus of external contemporary voices." What will be Polaroid's next big invention? Land laughs, but won't say. "I hope they guess what I hope it will be", he replies.

Deborah Shapley

CORRESPONDENCE

Coal in reserve

SIR — Robert Olby's article "Britain's resources of coal and spent uranium fuel" (*Nature* 29 April; 296 797-801) puts into sharp perspective the opposing arguments about fossil fuels versus nuclear fuels.

For the first time, the previously murky distinctions between the different classes of coal reserves have been made fairly clear. Previous attempts to do this have usually been obscured by the question of the future of coal versus other forms of energy. For example, those who believe that limited government resources should not be invested in the coal industry have tended to use the most restricting definitions of coal reserves, coupled with the lowest figures of percentage extraction, in order to suggest that coal has no future. As Olby points out, referring to A.M. Clarke, the future of coal is hardly unaffected by government policy. However, Olby himself is impartial in the matter of policy; he is simply clarifying the distinctions between different classes of coal reserves.

His brief account of Jevons's much misunderstood work of 1865 is also impressive. Jevons's point is just as valid today and, apparently, just as misunderstood. The finite extent of our recoverable fuel reserves does not define the period of time for which we can exist as we are before we run out of fuel. It defines the *much shorter period* for which we can continue to expand for the same production-to-fuel consumption ratio. A statement such as "At present rates of consumption, we have enough coal for 300 years" is thus meaningless.

Mr A.M. Clarke, ex-Chief Geologist of the National Coal Board, has probably done more than anyone else in Britain to stir up thought on these issues. It is good to see that his work is acknowledged in your pages, despite the continuing efforts both inside and outside the National Coal Board to denigrate his achievements.

ANTON ZIOLKOWSKI

Department of Mining Engineering,
Delft University of Technology,
Delft, The Netherlands

Ball of fire?

SIR — On Tuesday, 3 August, shortly after 4.00 p.m., the Cavendish Laboratory and the surroundings were struck by lightning several times during an exceptionally intense storm. No structural damage ensued, but immediately after one of the discharges a ball of light was seen by a number of observers. Their descriptions are not entirely consistent but certain features are agreed upon well enough to enable a broad description to be given. The discharge apparently responsible struck near the centre of the Bragg Building, which runs east-west. An observer on the ground floor of the Mott Building, whose back was to the window, saw his room momentarily lit as if by a very bright object moving past rapidly towards the west, between the Bragg and Mott buildings. Another observer on the first floor saw the space between the buildings filled with a luminous haze at least to the first floor level, and on looking to the west noticed a blue-white light that he thought at first was a warning light on a distant tower. He

apparently noticed no motion, but his companion in the same room must have seen it an instant earlier for she had the impression that it was moving past and away, and possibly expanding as it went, being about the size of a grapefruit when first seen. Three people who saw it after that, as it moved over the ground to the west, agreed it looked about the size of the moon, was blue-white in colour, very bright, and was visible for some 4-5 seconds before suddenly vanishing.

To this reasonably well attested observation must be added that while an assistant in the duplicating room, on the ground floor, was closing a small window she was startled by a noise that made her think the window had been knocked in; a bright sparkling object, resembling the lights thrown out by expensive rockets, entered by her head, rebounded from a machine and left as it came. The window was in fact undamaged, and when examined next morning entirely unmarked. Both assistants who were there at the time are convinced something came into the room.

BRIAN PIPPARD

Department of Physics,
University of Cambridge, UK

Badgers still snared

SIR — I understand that the Royal Society for Nature Conservation, while welcoming the ending of the gassing of badgers by the Ministry of Agriculture (*Nature* 22 July, p.317) because of its inhumanity, is adamant that only badgers that are positively found to be infected with tuberculosis should be destroyed. We learn, however, that snaring of badgers is to continue (which is surely as inhumane if not more so than their gassing) which makes it impossible to test for infection before killing.

GWENDOLEN BARTER

Margate, Kent, UK

Nuclear risks

SIR — I find the claims of Martin Fodor concerning the dangers of nuclear power (*Nature* 22 July, p.320) both confusing and inconsistent with the available evidence. Studies in a number of countries have compared the number of fatalities from the routine operations of the complete fuel cycle for nuclear, coal-fired and oil-fired generation. They all show that nuclear generation is at least as safe, and probably a good deal safer, per unit output than using coal or oil. Martin Fodor is quite wrong to imply otherwise.

These evaluations include the contribution from uranium mining to the hazards of nuclear generation. I am not aware of any studies showing 50-100 per cent occupational mortality among uranium miners, but it has been known since the fifteenth century that underground metalliferous mines give rise to enhanced mortality from lung disease among the workforce. Modern radiation safety standards based on the recommendations of the International Commission for Radiological Protection (ICRP) are designed to ensure that working with radiation safety ensures that the main occupational hazard in uranium mining is the physical risk.

Comparative estimates of the routine risks from the alternative energy sources are not so reliably established as for nuclear and fossil fuels. However, such information as is available suggests that the risk per unit generation is within the range of the other fuels. Professor Fremlin has pointed out that the idyllic village millpond has only to account for one drowning every 200 years to be classified as the most hazardous energy source, and generating electricity from wavepower, the renewable resource with which the United Kingdom is most generously endowed, is likely to involve substantial risks.

Rare major accidents invariably give rise to a much smaller death rate than the routine operational risks for all technologies. Despite this, they are of particular concern because of society's abhorrence of multiple fatality accidents. Nuclear electricity generation is particularly well protected against such events: not a single death from radiation has been recorded from any accident during the 25 years such plants have been operating. Whilst the probability of nuclear power accidents is certainly not zero, all the evidence indicates that the risk is much smaller than for many other accepted activities. Anyone concerned about accidents which may kill large numbers of people should worry about technologies other than nuclear power which are over 100 times more likely to cause such accidents.

The overall conclusion can only be that no practicable method of electricity generation is particularly hazardous, nor is any so safe as to merit particular preference. Nuclear electricity generation is certainly not the most hazardous of the technologies we use, and may be one of the safer options.

Safety is only one aspect of the environmental impact of electricity generation. In other respects nuclear power has an inherent advantage because it is the most compact energy resource available. Consideration of land use, visual intrusion, water supply, transport requirements and biological effects, together with the safety aspects, gives good grounds for the view that on balance nuclear power is the least environmentally damaging of the available options for providing the electricity we need.

BRIAN WADE

AERE Harwell,
Oxfordshire, UK

Oxygen treatment

SIR — In view of the recent press publicity, which we understand has embarrassed many neurologists and general practitioners, on the use of hyperbaric oxygen in the treatment of multiple sclerosis, the Multiple Sclerosis Society wishes to state that at no time has it recommended its members to seek treatment with hyperbaric oxygen.

We are awaiting with interest the results of a carefully monitored double-blind controlled study recently undertaken in America with the support of the American National Multiple Sclerosis Society. Upon receipt of the results of this trial the society's medical advisers will immediately review this matter.

JOHN WALFORD

Multiple Sclerosis Society,
London, UK

NEWS AND VIEWS

Mitochondrial mosaics —
maturases on the move

from Leslie A. Grivell and Piet Borst

To the initiated, the idea that mitochondria have opted for unusual modes of gene expression must surely be becoming a familiar one. A paper just published in *Nature* (298, 628; 1982) by Dujardin *et al.* shows, however, that additional surprises may still be in store regarding the proteins that mtDNA encodes. The paper illustrates that mutations are not necessarily changes for the worse: a single amino acid substitution in an intron-encoded protein is responsible for generating a new enzymatic function, leading to the expression of a novel mRNA maturase active in splicing.

Two years ago, Lazowska *et al.*¹ showed that a protein encoded in part by the second intron in the yeast mitochondrial gene for cytochrome *b* (*cob*) is required for the splicing out of that same intron sequence from its own mRNA, one of the precursors to the mRNA for cytochrome *b*. The protein, *box3* (or *bl₄*) RNA maturase, is a chimaera, consisting of the amino-terminal half of cytochrome *b*, fused to an intron-specific sequence. Since it controls its own synthesis by destroying the mRNA which encoded it, Lazowska *et al.*¹ coined the term 'splicing homeostasis' to describe the self-regulation of RNA processing that results.

Other mitochondrial introns, both in yeast and other fungi, also contain long unassigned reading frames (URFs)²⁻⁵ and many, if not all of them, may encode proteins with maturase action. Two series of observations support this view. First, the URFs display varying degrees of sequence conservation, suggestive of constraints on protein structure and indicating that their products have similar functions^{3,6,7}. Second, *trans*-recessive mutants defective in splicing have been identified in the URFs of the other two main introns of the cytochrome *b* gene (*bl₃* and *bl₄*)³⁻¹⁰.

An unexpected complication was the finding that numerous intron mutations of the cytochrome *b* gene are pleiotropic. They prevent the synthesis not only of cytochrome *b*, but also of subunit I of cytochrome *c* oxidase¹¹. This protein is also specified by a mosaic gene (*oxi3*) and the lack of expression is known to be due to a defect in the processing of the oxidase pre-mRNA^{12,13}. The effect on expression works in only one direction as mutations in *oxi3*, including large deletions, have no effect on the expression of the cytochrome

b gene^{12,13}. The link is positive as deletion of all or part of the cytochrome *b* gene similarly blocks oxidase synthesis¹⁴. The key factor has been traced by genetic experiments to the fourth intron of the long form of the cytochrome *b* gene (*bl₄*)¹⁵. Several lines of evidence^{9,16} now strongly suggest that the RNA maturase encoded by *bl₄* is also involved in excision of *al₄*, the fourth intron of the *oxi3* gene and one which is highly homologous to *bl₄* (ref. 3).

To learn more about the requirements for mitochondrial splicing and about the components involved, Slonimski and his associates have developed a general method for the selection and characterization of extragenetic suppressors of splicing-deficient mutations located in introns¹⁷.

One of these mutations has been fully described in the recent report by Dujardin *et al.* The mutation (*mim* [mitochondrial-mitochondrial interaction] 2-1) is a single G→A transition in the open reading frame of *al₄*, where it produces a glutamate to lysine substitution. In its presence, the cytochrome *b* gene and the mRNA maturase encoded by *bl₄* are no longer required for the splicing and expression of the oxidase gene and, just as remarkable, expression of the cytochrome *b* gene in mutants with a defective *bl₄* maturase is once again possible. Evidently, the *mim* 2-1 mutation induces the synthesis of an *al₄* maturase with an activity different from that present in wild type. This novel maturase can presumably act as a substitute for a defective *bl₄* maturase, using the same set of recognition sequences present in both pre-mRNAs¹⁶. This explanation is appealingly simple; nevertheless it raises a number of intriguing questions about the origin, evolution and function of the URFs in yeast mtDNA, some of which are addressed by Dujardin and his colleagues.

First, how is a new maturase formed from an 'inactive' one and what is the normal function of the *al₄* URF? Dujardin *et al.* propose that the replacement of a glutamate by lysine, which changes the overall charge of the protein by +2, may bring about a local increase in positive

charge in a region of the protein that is essential for catalytic activity. The *mim* 2-1 change occurs in a α -helical region of 15–18 amino acids, which occupies homologous positions in the *bl₄* and *al₄* URFs. The α -helix of wild-type *bl₄* is more basic than that of wild-type *al₄* by two or four charges, dependent on strain polymorphism. The *mim* 2-1 mutation would make the *al₄* translation product more similar to its *bl₄* counterpart.

What is the normal function of the *al₄* translation product? The *al₄* and *bl₄* URFs display a high (>70 per cent) homology at the protein level and are presumably derived from a common ancestor. On this basis Dujardin *et al.* consider two scenarios. First, the *al₄* protein may simply be a pseudogene. A single substitution might (somewhat implausibly) still be capable of resurrecting the gene and restoring its function. Second, evolutionary divergence of *al₄* and *bl₄* could have led to a diffusible protein, endowed with a pleiotropic splicing activity (*bl₄*) and to a protein with no catalytic splicing activity at all (*al₄*). A selective pressure against the accumulation of chain-termination mutations in *al₄* would be provided if it had retained, or acquired, other functions, for example as a structural component of the mitochondrial splicing and/or translation apparatus. However, because of the extent and nature of sequence conservation between the two URFs, we prefer the idea that the wild-type *al₄* translation product fulfils a maturase function, that has so far escaped detection by mutagenesis. The main effect of the *mim* 2-1 mutation may then be to alter substrate specificity.

What do maturases do and how do they do it? These are somewhat embarrassing questions, since not a single maturase has been isolated. They appear to be present in such low concentrations in wild-type yeast cells that they even escape detection by pulse-labelling. Conclusions about their structure and mode of action can, therefore, only be drawn from their DNA sequences and other indirect evidence. All URFs code for proteins which are rich in basic and hydrophobic residues and whose overall characters are suggestive of globular, non-membrane-associated proteins. Characteristics seen in similar distributions of charged residues, or residues

Leslie A. Grivell and Piet Borst are in the section for Medical Enzymology and Molecular Biology of the Laboratory of Biochemistry, University of Amsterdam.

disrupting helical structure^{6,7}, are conserved within a central 115 amino acid stretch of each URF. This region, which is bounded by characteristic decapeptide motifs, is thus likely to play a key part in determining both protein structure and enzymatic function, as does a region immediately downstream, in which numerous *trans*-recessive mutations affecting splicing have been identified^{9,10}.

Two indirect observations concerning maturase function are important. First, intron mutants, with defective mitochondrial maturases, generally accumulate only high molecular weight, intron-containing pre-mRNAs^{12,13}. Second, nuclear mutants defective in mitochondrial splicing reactions have been isolated and partially characterized. The defects are surprisingly specific, often involving an inability to cut or ligate the ends of a single intron in the gene for cytochrome *b* or subunit I of cytochrome *c* oxidase. In one group of cytochrome *b*-deficient mutants recently studied by Diekmann *et al.*¹⁸, transcripts consisting of bI₄ linked to downstream exon sequences accumulated, indicating first that a nuclear gene product is required for scission of the bI₄-exon boundary and second, by implication, that the bI₄ maturase may itself be involved in the initial scission at the upstream exon-bI₄ border.

With such an intricate systems for RNA processing, it is remarkable that many individual introns and their maturases are dispensable, in that various yeast strains, either found in nature¹⁹ or constructed in the laboratory^{8,15}, can lack them entirely and suffer no ill-effect. So, why have introns?

One view is that they are evolutionary relics, left over from a primitive, genes-in-pieces type of genome organization. Indeed, the high homology observed between introns in mtDNA of yeast and *Aspergillus nidulans*^{4,5} indicates that at least some introns pre-date the common ancestor of these fungi. On the other hand, whatever role introns may have had in the evolution of mitochondrial genes, their continued retention in fungal mtDNAs cannot be an accident. The distribution of introns is highly non-random: two of the seven major protein-coding genes, one of the two rRNA genes and none of the tRNA genes contain introns. The maintenance of the introns requires a large number of genes involved in splicing; this complex system must be vulnerable to disruption by mutation. Nature is not without means to eliminate this costly genetic investment, because it can produce mtDNAs with fewer or no introns. The introns in the fungal mtDNAs must therefore confer a selective advantage now.

Among possible advantages, the most plausible is that additional splicing steps allow a more precise control over mitochondrial gene expression. This could permit, for example, the selective suppression of the synthesis of mitochondrial

components during growth in the presence of fermentable sugars. Such an explanation may not, however, hold for all introns and all fungi. The gene for the mitochondrial large rRNA contains an intron plus URF, yet accurate splicing is possible in the absence of the URF product²⁰ (albeit somewhat inefficiently) and rRNA synthesis does not seem to be selectively repressible. Again, *Aspergillus* has also retained introns and their reading frames in the genes for cytochrome *b* and subunit I of cytochrome *c* oxidase, yet does not repress mitochondrial synthesis. There is obviously room for further speculation about the selective advantage of having mitochondrial introns and maturases.

Finally, where do the introns and their

maturases come from? We suggested previously²¹ that they might have arisen from insertion elements which lost their mobility in the course of evolution and this idea remains plausible. First, the sequence homologies found between the URFs of yeast mtDNA argue for a common origin with subsequent dispersal around the genome⁷. Second, the hypothesis correctly predicts that sequences homologous to intronic reading frames may also be found outside genes²². Third, all yeast URFs display codon usage which deviates markedly from the characteristic pattern observed in genes for known, major translation products^{3,23}. Such usage could be the last visible trace of a separate genetic origin of URF sequences. □

- Lazowska, J., Jacq, C. & Slonimski, P.P. *Cell* **22**, 333 (1980).
- Nobrega, F.G. & Tzagoloff, A. *J. biol. Chem.* **255**, 9828 (1980).
- Bonitz, S.G., Coruzzi, G., Thalenfeld, B.E., Tzagoloff, A. & Macino, G. *J. biol. Chem.* **255**, 11927 (1980).
- Waring, R.B. *et al.* *Cell* **27**, 4 (1981).
- Netzer, R., Köchel, H.G., Basak, N. & Kuntzel, H. *Nucleic Acids Res.* (in the press).
- Michel, F., Jacquire, A. & Dujon, B. *Biochimie* (in the press).
- Hensgens, L.A.M., Bonen, L., de Haan, M., Van der Horst, G. & Grivell, L.A. *Cell* (submitted).
- Jacq, C. *et al.* in *Mitochondrial Genes* (eds Slonimski, P.P., Borst, P. & Attardi, G.) 155 (Cold Spring Harbour Laboratory, New York, 1982).
- De la Salle, H., Jacq, C. & Slonimski, P.P. *Cell* **28**, 721 (1982).
- Weiss-Brummer, B., Rödel, G., Schweyen, R.J. & Kaudewitz, F. *Cell* **29**, 527 (1982).
- Pajot, P., Wambier-Kluppel, M.L., Kotylak, Z. & Slonimski, P.P. in *Genetics and Biogenesis of Chloroplasts and Mitochondria* (eds Bücher *et al.*) 443 (Elsevier, Amsterdam, 1976).
- Church, G.M., Slonimski, P.P. & Gilbert, W. *Cell* **18**, 1209 (1979).
- Van Ommen, G.J.B. *et al.* *Cell* **20**, 173 (1980).
- Alexander, N.J., Perlmann, P.S., Hansen, D.K. & Mahler, H.R. *Cell* **20**, 199 (1980).
- Dhawale, S., Hansen, D.K., Alexander, N.J., Periman, P.S. & Mahler, H.R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1778 (1981).
- Netter, P., Jacq, C., Carignani, G. & Slonimski, P.P. *Cell* **28**, 733 (1982).
- Dujardin, G., Pajot, P., Groudinsky, O. & Slonimski, P.P. *Molec. gen. Genet.* **179**, 469 (1980).
- Diekmann, C.L., Pape, I.K. & Tzagoloff, A. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1805 (1982).
- Borst, P. in *31. Colloquium-Mosbach 1980: Biological Chemistry of Organelle Formation* (eds Bücher, T., Sebald, W. & Weiss, H.) 27 (Springer, Berlin, 1980).
- Tabak, H.F. *et al.* *Nucleic Acids Res.* **9**, 4475 (1981).
- Borst, P. & Grivell, L.A. *Nature* **289**, 439 (1981).
- Coruzzi, G., Bonitz, S.G., Thalenfeld, B.E. & Tzagoloff, A. *J. biol. Chem.* **256**, 12780 (1981).
- Dujon, B. in *Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance* (Cold Spring Harbour Laboratory, New York, 1981).
- Hudspeth, M.E.S., Ainley, W.M., Shumard, D.S., Butow, R.A. & Grossman, L.I. *Cell* (in the press).

Where has that Moon been?

from Kurt Lambeck

THAT tidal forces play a most significant part in the dynamical evolution of the Earth-Moon system is beyond dispute. What is in dispute is the time scale of this evolution. One more attempt at resolving this vexing question has recently been discussed by Webb in the *Geophysical Journal of the Royal Astronomical Society* (**70**, 261; 1982).

That this ancient subject continues to draw the attention of scientists in diverse disciplines may seem curious but is a reflection of the important ramifications of tidal evolution models as well as their intellectually satisfying nature. Few other subjects require one to delve into the rigours of celestial mechanics or into the ambiguity of an observational record that includes eighteenth century telescope observations, ancient records of eclipses and the life styles of Palaeozoic corals. It requires investigations into the solid Earth's structure, into ocean tides, speculations on the evolution of the Earth and an initiation into the mysteries of lunar chemistry.

A central difficulty in modelling the past orbital and spin configurations of the

Moon and Earth is to extrapolate from present-day tide models and tidal accelerations. It is inescapable that most of the tidal energy is dissipated in the oceans¹ and that this must have been so for much of the Earth's past history. Any extrapolation backwards into geological time is, therefore, only as good as the assumption that past and present tides are identical or that past tides can be characterized by some mathematically convenient law. Such extrapolations lead to the conclusion that the Moon must have been very close to the Earth some 1.5–2 billion years ago. This conclusion is unsavoury to most lunar scientists².

Can this close approach, so late in the Earth's history, be avoided without invalidating the orbital evolution calculations? Is it possible to push it back into the Earth's remote past? This is the problem that Webb sets out to solve; how to compute representative tidal dissipation estimates for past epochs when the requisite detailed oceanic data are

Kurt Lambeck is Visiting Professor at the University of Paris-Sud, on leave from the Australian National University, Canberra.

unavailable. His approach is not to consider detailed models at specific geological epochs as others have attempted³. Instead Webb adopts an 'average' ocean model and then investigates how dissipation may vary in response to the evolving Earth-Moon system. In particular, he investigates the dependence of dissipation on the frequency of the tide-raising force. His important result is that the rate of tidal energy dissipation has been less in the past than it is now and that the calamitous consequences of a close approach of the Moon can be readily averted.

Present rates of energy dissipation are now quite well established as about 4.0×10^{12} Watts (ref. 4). It is also clear that at least 90 per cent of this energy is dissipated in the oceans. What remains less obvious is the mechanism by which this dissipation takes place. Is it by bottom friction in a few shallow seas, or by a more global mechanism such as the breaking of tidal waves along shorelines or the scattering of tidal waves into internal modes? Advocates of all mechanisms can be found but it seems that the actual assumptions made have little influence on the numerical solutions or interpolations of the Laplace tidal equations. Compare, for example, the recent tidal models for the dominant lunar tide⁵⁻⁸. All solutions require that dissipation is fairly uniformly distributed through the world's oceans, and all give very comparable results once corrections for the solid tide are made where appropriate⁴. That this is so is at least partly because all solutions are either explicitly or implicitly constrained by observations of the tides. But it makes Webb's approach of considerable interest for he does not have to make assumptions about the detailed ocean geometry. Instead, he considers tides on a hemispherical ocean⁹ whose orientation with respect to the Earth's rotation axis is allowed to vary through all possible values. Dissipation is introduced by requiring that the decay time of tidal energy equals the average observed value. For a given frequency of the tide — defined by the Earth-Moon distance and the Earth's spin rate — Webb averages the power estimates over all orientation angles with the acceptable rationale that the time constant for plate tectonics is much less than the age of the oceans. The variation in the ocean geometry with time is therefore accounted for in this statistical manner. At any one frequency, strong resonances are encountered as a function of the orientation but Webb finds that, when averaged, these isolated resonances are smoothed out and that any resonances in the averaged power curve occur at frequencies that are lower than present and past values: the further one goes into the past the greater becomes the separation between the forcing and resonance frequencies.

Consequently, past rates of orbital evolution will be less than if the Q or

dissipation function of the ocean had been held constant, and Webb concludes that the close approach may never have occurred at all. Quantitatively his results may not be very significant but the general trend that emerges is. Evidence for reduced dissipation in the geological past has also been found in palaeontological records of the Earth's rotation where growth rhythms in corals and bivalves point to less dissipation over the past 400 million years than occurs at present⁴. A similar conclusion was also reached from explicit model studies of past ocean tides³.

A word of caution seems appropriate however. The resonance frequencies of the ocean basins depend on the ocean scale as well as on the ocean depth. Thus by fixing the decay time, Webb also assumes that these factors have not changed much. The approximate period of the free oscillation is given by $2L/(gh)^{1/2}$ where L is the characteristic length scale, h the average depth and g is gravity. If, on the average, L is greater than today, as when the oceans

were clustered into the Pangea continent, then the driving force will have been further away from resonance than it is now. If the ocean depths were less in the past, if the present basins evolved only relatively late in the Earth's history, then the difference between the forcing and resonance frequencies also increases. Thus some cancellation of Webb's result may occur. I rather doubt that this will be significant and those who argue that the Moon formed close to the Earth can sleep in peace. □

1. Lambeck, K. *Phil. Trans. R. Soc. A* **287**, 545 (1977).
2. Ringwood, A.E. *The Origin of the Earth and Moon* (Springer, Berlin, 1979).
3. Sundermann, J. & Brosche, P. *Tidal Friction and the Earth's Rotation*, 125 (Springer, Berlin, 1978).
4. Lambeck, K. *The Earth's Variable Rotation* (Cambridge University Press, 1980).
5. Zahel, W. *Ann. Geophys.* **33**, 31 (1977).
6. Accad, Y. & Pekeris, C.L. *Phil. Trans. R. Soc. A* **290**, 235 (1978).
7. Parke, M.E. & Hendershott, M.C. *J. mar. Geod.* (1980).
8. Schwiderski, E.W. *Rev. Geophys. Space Phys.* **18**, 243 (1980).
9. Webb, D.J. *Geophys. J. R. astr. Soc.* **61**, 573 (1980); **68**, 689 (1982).

From CoA to complement: thioesters as the spring in the molecular mouse trap

from John Fothergill

To most biochemists, thioesters mean coenzyme A and fatty acid metabolism. They are remarkable for their high free energy of hydrolysis and their participation as metabolic intermediates. With this background, the discovery of thioesters in proteins came as something of a surprise.

There is now good evidence that two of the complement proteins, C3 and C4, and the plasma proteinase inhibitor α_2 -macroglobulin contain an internal thioester formed from the side chains of a cysteinyl and a glutamyl residue four positions apart on the polypeptide chain. In the native protein, the thioester is apparently buried, but on activation of the protein by proteolytic cleavage the thioester is exposed to the surroundings, readily generating a free thiol and an active acyl group that can react with nucleophilic amino or hydroxyl groups. This leads to covalent attachment of the thioester protein to adjacent protein or carbohydrate. Thus a single proteolytic cleavage of the protein containing the thioester leads to covalent linking of the activated protein to a nucleophilic group in its immediate environment. What are we to make of this 'molecular mouse trap'?

It has been known for a long time that complement components, once activated, adhere strongly to cell surfaces. This was

explained as an interaction between a hydrophobic region on the complement component, revealed during proteolytic activation, and the cell membrane. The relatively short half life of the activated complement components was explained in terms of conformational rearrangement of this hydrophobic binding region. The first evidence¹ that the reaction could be a covalent one came from the detection of proteins in SDS-polyacrylamide gel electrophoresis that had higher molecular weights than the C3 polypeptide chains they were known to contain. Similar evidence has been obtained for the covalent binding of C4 (ref. 2) and C3 (ref. 3) to antigen-antibody aggregates. During the last two years evidence for this internal thioester has been accumulating rapidly. Publications have been overwhelmingly confirmatory, if somewhat repetitious.

Much of the evidence for the thioester in C3 has come from the work of Janatova

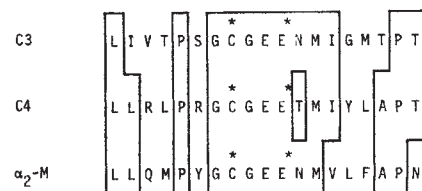


Fig. 1. Amino acid sequences around the thioester site of C3 (ref. 4), C4 (ref. 5) and α_2 -macroglobulin (ref. 6). The cysteinyl and glutamyl residues contributing to the thioester are marked by an asterisk.

John Fothergill is in the Department of Biochemistry, University of Aberdeen, Aberdeen AB9 1AS.

and Tack, and a recent paper⁴ summarizes the present position. The glutamyl residue of the thioester can be labelled with radioactive methylamine producing the corresponding γ -glutamylmethylamide; the thiol group reacts with the usual reagents. Labelling studies have shown clearly that the reactive residues are in the C3d part of the C3b molecule which is known to contain the membrane-binding site. The amino acid sequence of this part of the polypeptide chain has been determined and it is obviously homologous not only to the corresponding region in C4 (ref. 5) but also to the thioester sequence in α_2 -macroglobulin⁶ (Fig. 1). Model-building studies have shown that the thioester cannot form if the polypeptide adopts an α -helix, β -sheet or β -turn conformation, but it is possible to arrange the atoms without undue strain in an unusually sharp bend that allows the glutamyl carboxyl and cysteinyl thiol groups to approach close enough to form the thioester.

The whole thioester story is intimately bound up with the question of conformational rearrangement during activation. Intact C3 is known to show circular dichroism spectra that are significantly different from those of a mixture of its activation products C3a and C3b (ref. 7). Treatment of C3 with methylamine or by freezing and thawing also produces evidence of conformational changes⁸ and renders the C3 haemolytically inactive. Interestingly, evidence for an increase in exposed hydrophobic surface occurring concomitantly with this conformational change was also obtained⁹. Whether the biological activity of C3 and C4 depends on both covalent attachment by virtue of the thioester cleavage and hydrophobic interactions resulting from conformational rearrangement remains to be unequivocally demonstrated. It is noteworthy that the structurally homologous complement protein C5 has similar properties but no thioester.

The discovery of these thioesters has opened up many new possibilities in mechanisms of complement activation and in more general aspects of protein chemistry. The suggestion has been made that spontaneous hydrolysis of the thioester in C3 may generate C3b-like activity that could form the nucleus of a C3

convertase for the alternative pathway¹⁰. The implications for biosynthesis are particularly fascinating. At what stage of maturation is the thioester formed? Is there a thiokinase, with requirements for a

nucleoside triphosphate, or does the folding of the nascent polypeptide chain contribute all that is necessary? Is the Lippman 'squiggle' the spring of this molecular mouse trap? □

Focus on filaments: embryology to pathology

from Birgitte Lane and Brian Anderton

A RECENT workshop* provided a useful summary of current research on intermediate filaments. With actin and tubulin, these filaments constitute the third fibre system of the cytoskeleton. They are morphologically well defined (as 10 nm filaments) and stable *in vitro*, and show a biochemical heterogeneity which is developmentally regulated. Each class of intermediate filament is restricted *in vivo* by tissue differentiation: vimentin to mesenchymal cells, desmin to muscle tissues, keratins to epithelia, glial fibrillary acidic protein (GFAP) to astroglia and neurofilament proteins to neurones. Given this apparently ideal experimental system, it is frustrating that so little is yet known about intermediate filament function.

However, progress is being made. Some synthesis is becoming possible regarding molecular structure, as Geisler (Max Planck Institute, Goettingen) showed in proposing a model for vimentin-like proteins, based on primary sequence data. The model molecule has three domains separated by enzyme-sensitive hinge regions: (1) a short, basic, folded head piece (about 7.5K) sensitive to proteolysis; (2) a long (38K) middle piece, helical and very acidic; and (3) a short tail piece, about 5.5 K and probably globular. The middle piece helix has a hydrophobic edge which may stabilize the intertwining of molecules into protofilaments, while charged patches on other parts of the helix may aid association into 10 nm filaments. Mid-pieces can aggregate into filamentous and sheet-like structures, but the head and tail pieces seem to be necessary for correct assembly of filaments.

Quax (University of Nijmegen) reported the DNA sequence of a cloned hamster vimentin gene: there may only be a single gene for vimentin in mammals, in contrast to the multiple genes for keratins described by Green (Harvard Medical School). Surprisingly, keratin, and not vimentin, is the first type of intermediate filament to appear in ontogeny. The first keratins appear in the morula (Jackson, University of Geneva) as a pair of polypeptides of 53K (basic group) and 44K (acidic group). Franke (German Cancer Research Centre,

Heidelberg) declared this to be the simplest keratin profile, therefore presumably the minimum required for filament formation. It was clear from many two-dimensional gels shown at the meeting that keratins can be classified as either acidic (smaller) or basic (generally larger). Presumably it is this complementarity of charge and size that restricts the major epidermal keratins to obligate heteropolymerization, as described by Steinert and colleagues some years ago.

Although some adult epithelia retain an almost embryonic simplicity of keratins, most differentiating epithelia acquire a more elaborate keratin profile, culminating in epidermal keratinocytes. Franke reported that 19 separate human keratins have now been identified. Sun (Johns Hopkins Medical Institute) suggested an ontogenic 'tree' model to help visualize the developmental complexity of the epidermis, which he demonstrated using three complementary monoclonal antibodies. Monoclonal antibodies to keratins can be used to detect non-keratinocytes in complex epithelial tissues (Lane, ICRF), and an elegant presentation by Moll (German Cancer Research Centre) showed how keratin microheterogeneity within a tissue or tumour could be scrutinized by two-dimensional gel electrophoresis of laser-dissected frozen sections.

Co-expression *in situ* of different types of intermediate filaments in one cell has only been reported in restricted situations, usually involving an overlap early in development between vimentin and the later-appearing tissue-specific class. Two cases of filament co-expression in adult cells were described at the meeting: astroglia, which by consensus of opinion usually contain vimentin as well as GFAP, and aortic smooth muscle, where cells may contain desmin or vimentin or both (Frank, Wistar Institute). Otherwise, the tissue specificity *in situ* of the non-vimentin filament classes remains clear-cut.

Do tumours express the filament type characteristic of the normal cell from

Birgitte Lane is at the Imperial Cancer Research Fund Laboratories, Lincoln's Inn Fields, London WC2A 2PX, and Brian Anderton is at St George's Hospital Medical School, Cranmer Terrace, London SW17 0RE.

1. Law, S. K. & Levine, R. P. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2701 (1977).
2. Campbell, R. D., Dodds, A. W. & Porter, R. R. *Biochem. J.* **189**, 67 (1980).
3. Gadd, K. J. & Reid, K. B. M. *Biochem. J.* **195**, 471 (1981).
4. Thomas, M. L., Janatova, J., Gray, W. R. & Tack, B. F. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1054, (1982).
5. Campbell, R. D., Gagnon, J. & Porter, R. R. *Biochem. J.* **199**, 359 (1981).
6. Swenson, R. P. & Howard, J. B. *J. biol. Chem.* **255**, 8087 (1980).
7. Hugli, T. E., Morgan, W. T. & Müller-Eberhard, H. J. *J. biol. Chem.* **250**, 1479 (1975).
8. Isenman, D. E., Kells, D. I. C., Cooper, N. R., Müller-Eberhard, H. J. & Pangburn, M. K. *Biochemistry* **20**, 4458 (1981).
9. Isenman, D. E. & Cooper, N. R. *Molec. Immun.* **18**, 331 (1981).
10. Pangburn, M. K., Schreiber, R. D. & Müller-Eberhard, H. J. *J. exp. Med.* **154**, 856 (1981).

*The EMBO-sponsored workshop 'Intermediate filaments in differentiation and pathology' was held at the end of April in Gunzburg, FRG.

which they arise? The answer to this question is clearly yes — tissue specificity of intermediate filaments is retained in the tumour, and vimentin does not appear inappropriately. This was demonstrated by Osborn (Max Planck Institute, Goettingen) in the keratin-positive, vimentin-negative phenotype of intestinal cancers, and the desmin-positive, vimentin-negative phenotype of rhabdomyosarcomas.

On the other hand, tissue culture can modify filament expression, in that vimentin appears in most cultured cells. The meeting heard the first report of vimentin in epithelial cells in the body, however. Ramaekers (Catholic University of Nijmegen) observed vimentin in carcinoma cells growing as ascites, that is, in suspension, but when growing as a solid tumour these cells only expressed keratin. This may be a clue to the significance of vimentin induction.

Pathologists at the meeting agreed that intermediate filament antibodies are useful for identifying the origin of tumour metastases when standard histological methods are inconclusive. Improvements were suggested to avoid the need for frozen sections, since most intermediate filament antibodies react poorly with aldehyde-fixed material. Osborn showed high-quality immunofluorescence preparations of methanol-fixed, paraffin wax-embedded material; Coggi (University of Milan) used formalin-fixed material, but amplified sensitivity with multiple biotin-avidin complexes.

It is sometimes of immense medical value to be able to identify the origin of displaced cells. Virtanen (University of Helsinki) reported the presence of GFAP-positive (that is, astroglial) cells in amniotic fluid from anencephalic fetuses, and in many cases of spina bifida; normals were negative, since unless nervous tissue is exposed the glial cells cannot get into the amniotic cavity.

Two examples of a pathological expression of intermediate filaments were discussed: Mallory bodies (containing keratin) of liver cells in alcoholic cirrhosis (Denk, University of Vienna) and neurofibrillary tangles of Alzheimer's senile dementia (Anderton, London, and Bignami, Harvard Medical School). Anderton reported that the neurofibrillary tangles could only be stained by some of a group of neurofilament-staining monoclonal antibodies, and similar partial cross-reactivity with conventional antisera was reported by Bignami, implying that neurofibrillary tangles contain structurally altered neurofilaments.

Perhaps the most unexpected reports came from studies of mitosis. Franke and Lane independently described how keratin filaments break up into balls during mitosis in epithelial cell lines. This fact is not widely known, probably because of the preference for fibroblast models in cytoskeleton studies. Lane described how the balls, which form in prophase, may be associated

with the plasma membrane, and Franke described associated 2 nm filaments, and said there was no evidence for the involvement of proteolysis. This disruption during mitosis is apparently not a tissue culture artefact, since Franke showed the effect in a human carcinoma. It was agreed that not all cell lines show the phenomenon, and Lane suggested a link with the speed of the cell cycle, such that more slowly dividing cells maintain intact filaments. Since keratins are difficult to

depolymerize in non-denaturing conditions, this may be an interesting experimental system.

Thus, with some useful data on molecular structure, and the availability of systems in which the cell reorganizes its filaments *in situ*, we may soon understand how, and why, these molecules interact in the cell. Meanwhile the usefulness of filament antibodies in medical diagnosis should attract a greater commitment to this field of cell biology. □

Do steroid receptors recognize DNA sequences?

from J. R. Tata

ON the twentieth anniversary of the Jensen two-step model of steroid hormonal regulation of gene expression¹ there is, at last, good evidence for one of its chief predictions. According to the model, steroid hormones interact in their **target** tissues with a high-affinity cytoplasmic receptor protein, and the interaction leads to a conformational change in the receptor such that the hormone-receptor complex is rapidly translocated into the nucleus. The receptor or receptor-hormone complex is then thought to interact with a variety of nuclear components in order to induce or modulate transcription of a few well specified genes. A prime nuclear target is DNA itself, but only recently has there emerged the first strong evidence to suggest sequence specificity of the interaction, at least *in vitro*.

In a recent study from Strasbourg, Mulvihill *et al.*¹ have most effectively exploited recombinant DNA technology and affinity chromatography to investigate the possible sequence specificity of interaction between the progesterone receptor and the genes of the egg-white protein that are induced by progesterone in the chicken oviduct. By adaptation of the oestrogen receptor DNA competition assay of Kallos and Hollander², Mulvihill *et al.* identified several cloned fragments from five chicken egg-white protein genes regulated by oestrogen and progesterone which could compete with calf thymus DNA for the progesterone receptor.

By testing deletion mutants, produced *in vitro*, of a fragment of the ovalbumin gene, Mulvihill *et al.* were able to narrow down the receptor recognition site to a region 250–300 base pairs (bp) 5'-upstream from the transcription initiation site, although recombinant fragments from other regions of all five genes (including exons and introns) were also found to be efficient com-

petitors for the receptor. They then derived a consensus sequence of 19 nucleotides



not detected in a similar region of the globin gene, which is thought to be recognized by the progesterone-receptor complex.

It is interesting to compare the results of the Strasbourg group with findings from Schrader and his co-workers in Houston who have used a direct nitrocellulose DNA-binding assay^{3–5}. Schrader also recently reported an A + T-rich region 135–250 bp upstream from the mRNA start site which is recognized by the progesterone receptor⁶. Photoaffinity labelling of the receptor with a synthetic analogue of progesterone (R5020) enabled this group to confirm their earlier suggestion that only one of the two subunits (A subunit) of the receptor recognizes DNA, while the other binds to chromatin (presumably to a protein component). However, the results of the two groups differ in some important respects. Whereas Mulvihill *et al.* find that the progesterone receptor specifically recognizes and binds to double-stranded DNA, the Houston group reported that the progesterone receptor acts as an unwinding or helix-destabilizing protein and has a preference for single-stranded DNA. The Strasbourg group ascribes this major inconsistency to several differences in methodology. They stress that they have used a crude cytoplasmic extract, in contrast to the highly purified preparation used in Schrader's laboratory, and that over-purification may have removed another component essential for the sequence-specific recognition of double-stranded DNA. They also point out that their assay measures the extent of competition by soluble DNA of progesterone-receptor complex to DNA already bound to cellulose, whereas the DNA filter-binding assay only gives the rate of association between the receptor and DNA.

J. R. Tata is Head of the Laboratory of Developmental Biochemistry at the National Institute for Medical Research, Mill Hill, London NW7 1AA.

The sequence specificity of DNA binding by another steroid hormone receptor has also been reported recently. In a collaborative study between the University of California, San Francisco, and the Karolinska Institute in Stockholm, Payvar *et al.*⁷ have suggested that the receptor for glucocorticoid hormones recognizes specific DNA sequences of murine mammary tumour virus (MMTV) within or neighbouring the gene, although the precise location of all possible sequences has not yet been determined. Their receptor preparation was purified 10,000-fold (that is, to about 50 per cent homogeneity) and tagged with the synthetic glucocorticoid ³H-triamcinolone acetonide as the radioactive ligand. In a separate report, the San Francisco group has described how a 4.5 kilobase pair (kbp) *EcoRI* fragment of MMTV DNA cloned in plasmid pMTV2, which contains the long terminal repeats encoding the transcription start site but lacks 4.4 kbp of the rest of MMTV DNA, induced glucocorticoid-responsive transcription when introduced into cellular DNA⁸. Payvar *et al.* showed by ³²P end-labelling that the MMTV DNA fragment was specifically retained on nitrocellulose filters when bound to the glucocorticoid receptor preparation. Three different types of competition binding experiments demonstrated that the receptor selectively binds to at least four widely separated regions of MMTV proviral DNA. One of these binding sites was located 110–449 bp upstream of the promoter for MMTV transcription. Preliminary observations (for which experimental data have not yet been provided) revealed that at least one receptor binding site is located several kilobase pairs downstream from the start site of MMTV transcription.

How relevant are these *in vitro* receptor DNA binding studies to the physiological action of steroid hormones? Pongs and his colleagues in Bochum have provided the first direct evidence for the *in vivo* localization of a steroid hormone–receptor complex at the site of hormone-activated genes⁹. In the first of two papers, they showed that the insect metamorphic hormone β -ecdysterone can be covalently fixed by photolabelling to specific hormone-induced puffs of the polytenic chromosomes in the salivary glands of the third instar larvae of *Drosophila*. By using the 'Western' blotting technique of localization of ecdysterone antibody on cellulose nitrate filters, and two-dimensional gel electrophoresis, the second paper convincingly demonstrates that UV irradiation cross-links ecdysterone to a single protein of molecular weight 130,000 in both *Drosophila* larval salivary glands and *Kc Drosophila* tissue culture cells. This single protein exhibits all the characteristics of a steroid hormone receptor, including translocation from the cytoplasm to the nucleus.

The technical innovations in these recent

studies represent a major step forwards in unravelling the role of steroid hormones in regulating gene expression, but they leave unanswered some important questions and raise a few new ones. First of all, discrepancies resulting from different techniques must be resolved. Then, there is the unresolved question of the high number of nuclear receptor molecules associated with full hormonal response. Although 42 sequences homologous to the 19 bp consensus segment were detected in the 14 kbp of known egg-white protein gene sequences, this number, even after generous extrapolation, does not explain the several thousand receptor binding sites normally found per cell nucleus. In this context, it is worth pointing out that the relative intensity of interaction between the progesterone receptor and specific 5'-upstream DNA sequence is only 10–40 times that observed with non-specific sequences such as calf thymus DNA or cloned globin genes^{1,5}. Does this fact suggest that numerous other binding sites, not on DNA but on chromosomal proteins (also known as 'acceptor' sites), ribonucleoproteins, nuclear matrix and even non-nuclear sites, are required to induce alterations in overall structural conformation in order fully to maintain transcriptional activity? Will the results

with progesterone receptor apply equally well to oestradiol, the major physiological regulator of egg-white gene expression?

In vivo functional tests, such as the introduction of viral sequences in a cell containing the rat glucocorticoid receptor, or the photoaffinity bonding of ecdysterone to its receptors in *Drosophila* cells, represent important approaches to be explored further. The next major advance will be the successful introduction of receptors in cells lacking these regulatory molecules and the manipulation of genes coding for the receptors. There is still some way to go before we can derive a unified concept of how hormones regulate gene activity, but what we learn along the way will be most valuable for our understanding of the multiple mechanisms of eukaryotic gene expression. □

1. Mulvihill *et al.* *Cell* **24**, 621 (1982).
2. Kallos & Hollander *Nature* **272**, 177 (1978).
3. Coty *et al.* *J. Steroid Biochem.* **10**, 1 (1979).
4. Hughes *et al.* *Biochemistry* **20**, 2481 (1981).
5. Compton *et al.* *Biochem. biophys. Res. Commun.* **105**, 96 (1982).
6. Schrader *UCLA Symp. on Gene Regulation*, Keystone, 28 March–3 April (1982).
7. Payvar *et al.* *Proc. natn. Acad. Sci. U.S.A.* **78**, 6628 (1981); *UCLA Symp. molec. cell. Biol.*, Squaw Valley (March 1982).
8. Yamamoto *et al.* *Cold Spring Harb. Symp. quant. Biol.* **45**, 681 (1981).
9. Gronemeyer & Pongs *Proc. natn. Acad. Sci. U.S.A.* **77**, 2108 (1980); Schaltmann & Pongs *Proc. natn. Acad. Sci. U.S.A.* **79**, 6 (1982).

Molecular heterogeneity and the nervous system

from Colin J. Barnstable

THE nervous system has an anatomical and physiological complexity that is considerably greater than that of other tissues and is seemingly part of a generally increased level of molecular complexity. The complexity arises both because the nervous system consists of more cell types and subtypes than other tissues and because a given neural cell seems to have a greater molecular complexity. Monoclonal antibodies are now providing an important new means of tackling this heterogeneity and of recognizing similarities and differences in neural cell types. Indeed, the use of monoclonal antibodies has grown so rapidly that two major review volumes on the fledgling subject of 'neuroimmunology' have already been published^{1,2}. Two recent papers^{3,4} provide good examples of the approach currently being taken by a number of laboratories.

In one paper, Hawkes *et al.*³ used a synaptic plasma membrane fraction from rat cerebellum to immunize mice for monoclonal antibody production. In the

other, Sternberger *et al.*⁴ used a crude homogenate of rat hypothalamus as the immunogen.

Both papers describe antibodies that stain sections of rat brain in a manner characteristic of neurofibrillar staining. This makes it likely, though not certain, that these antibodies are staining neurofilaments, the class of intermediate filaments unique to neurones. Unlike intermediate filament types of other tissues, which consist of similarly sized polypeptide chains, neurofilaments have been isolated as a triplet of distinct polypeptides of molecular weights 200,000, 140,000 and 68,000 (ref. 5). Hawkes *et al.* found one such antibody that stained the cerebellum and reacted with a polypeptide of molecular weight 30,000. It is not yet clear whether this represents a new neurofilament polypeptide or whether it is some neurofilament-associated molecule. The staining patterns given by the other neurofibrillar antibodies were not identical, which suggests that either the neurofilament polypeptides are distributed differently between various neuronal cell types or that one or more of the neurofilament polypeptides exists in a

Colin J. Barnstable is in the Department of Neurobiology, Harvard Medical School, Boston, Massachusetts 02115.

variety of isotopic forms. Biochemical analysis of the antigens recognized by these monoclonal antibodies should resolve this question.

An important point to emerge from these papers is that an antibody can label different subpopulations of neurones in different brain regions. An example given by Sternberger *et al.* is the antibody 02-135 which labels pyramidal cells in area CA2 but not CA1 of the hippocampus, Purkinje cells but not granule cells of the cerebellum, and stellate cells but not pyramidal cells of the cerebral cortex.

The interpretation of such results is complex. Similar results using antibodies raised against the leech nervous system were interpreted as suggesting a possible functional interaction or neural sub-network between the labelled cells⁶. While this is indeed possible, mammalian studies suggest a different interpretation. It seems as if each antigen can be used for one class of cells in an area of the nervous system that is physiologically and anatomically separated from other regions in which the same antigen appears. Clearly further characterization of the respective antigens is required since any interpretation will depend on whether or not the same molecule is being detected in the various cells. Hawkes *et al.* have shown that some antibodies that give very distinct immunocytochemical staining can react with a number of polypeptides of different molecular weights and the general problems of cross-reaction of monoclonal antibodies have been discussed in these pages recently⁷. If we view the nervous system as having evolved by duplication and divergence of cell types or even whole neuronal networks, then the problem of cross-reactions between related molecules is going to be particularly acute.

One antibody produced by Hawkes *et al.* has been characterized more fully⁸ and shown to recognize a 23,000 molecular weight polypeptide that is present in all neurones. Both anatomical studies and subcellular fractionation experiments suggest that this antigen, named MIT-23, is restricted to neuronal mitochondria. During cerebellar development this antigen appears after the final cell division or migration of the neurones. It is at present unknown whether this molecule carries out a function that only appears in postmitotic neurones or whether it is a neurone-specific isozyme of a common mitochondrial component, another isozyme of which is present in the

dividing neuroblast. This latter possibility would be reminiscent of one of the first immunologically identified 'brain-specific antigens', the 14-3-2 protein, that turned out to be a brain-specific subunit of the enzyme enolase that is only expressed in postmitotic neurones⁹.

Because of the different methods being used to characterize antibodies by different laboratories, and the different extents of characterization, it is often difficult to compare results in different papers. A

particularly acute problem arises when some of the results are based on reactions of uncloned antibodies, as in the case above. One thing that is clear, however, is that we do not yet know the full extent of the molecular complexity of the nervous system. The challenge facing those who would like to find molecules carrying out a particular function is to identify them against what is clearly a high background of molecular complexity that may be completely incidental to those functions.

BIRPS: deep seismic reflection profiling around the British Isles

from D.H. Matthews

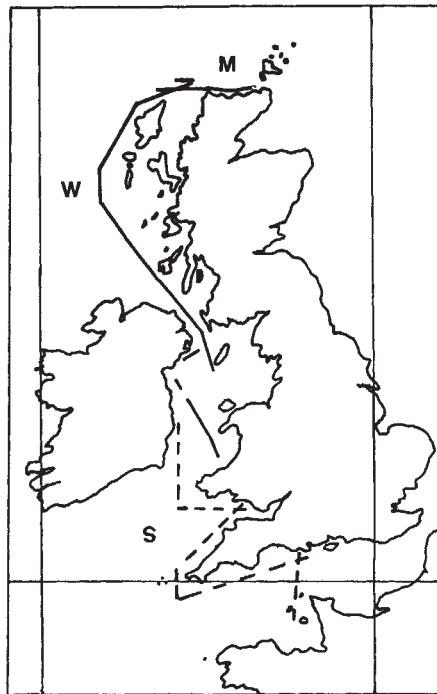
BIRPS* is investigating the geological structure of the lower part of the continental crust and of the underlying uppermost part of the Earth's mantle in and around the British Isles, using seismic reflection profiling techniques developed by the oil industry. The technique can be used at sea or on land and is similar to echo sounding: sound from an explosion at the surface passes down through the rock layers and is reflected back to a detector at the surface from inter-layer boundaries, across which there is a change in acoustic impedance (the product of sound velocity and density).

Two lines (MOIST and WINCH) have been shot so far — both offshore, as shown on the map (Fig. 1). The sound source was an array of air-guns which released compressed air into the sea every 50 m along the ship's track. Reflected sound was detected by a towed streamer, 3 km long, containing hydrophones arranged in 60 sections each 50 m long. By adding the returns from 30 consecutive shots recorded from sections 2, 4, 6 . . . 58, 60 and correcting for the increasing length of the raypath it was possible to add 30 reflections from a common point on the reflector (30-fold coverage). The recording equipment was run for 15 s; sufficient for the sound to travel to about 45 km beneath the surface and back again if the mean crustal sound velocity is 6 km s⁻¹.

MOIST obtained what are arguably the best deep seismic reflection data in the world^{1,2} (Fig. 2). The crust-mantle boundary, the Moho, can be traced as a strong reflector along the greater part of

the profile westwards from the Pentland Firth until it apparently terminates against a second strong reflector (the Flannan Thrust?) which can be followed down to 40 km depth. The Moho depth of 27 km agrees with that obtained on the LISPB refraction line where the two lines intersect³. Further velocity information was obtained from two-ship multichannel experiments⁴. The Outer Isles Thrust, which was reactivated as a normal fault in Mesozoic time, can be traced down as a planar reflector dipping at about 30°

Fig. 1 Outline of the British Isles with the positions of deep reflection lines: MOIST (Moine and Outer Isles Seismic Traverse, 180 km, shot 1981 by Western Geophysical), WINCH (Western Isles and North Channel Traverse, 1,000 km, shot 1982 by GECO International), SWAT (South West Approaches Traverse, planned for 1983). Meridians 0° and 10°W and parallels 50° and 60°N (the top border) are shown.



- McKay, R., Raff, M. C. & Reichert, L.F. (eds) *Monoclonal Antibodies to Neural Antigens* (Cold Spring Harbor Laboratory, New York, 1981).
- Brookes, J. (ed.) *Neuroimmunology* (Plenum, New York, 1982).
- Hawkes, R., Niday, E. & Matus, A. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2410 (1982).
- Sternberger, L.A., Harwell, L.W. & Sternberger, N.H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1326 (1982).
- Lazarides, E. *Nature* **283**, 249 (1980).
- Zipser, B. & McKay, R. *Nature* **289**, 549 (1981).
- Lanc, D. & Koprowski, H. *Nature* **296**, 200 (1982).
- Hawkes, R., Niday, E. & Matus, A. *Cell* **28**, 253, (1982).
- Bock, E. *J. Neurochem.* **30**, 7 (1978).

*BIRPS (The British Institutions Reflection Profiling Syndicate) is made up of members of university earth science departments and the Institute of Geological Sciences. The core group receives an annual grant from the Natural Environment Research Council and Shell (UK) supports a visiting scientist. Data from the first seismic reflection line shot (MOIST) are on open file at the Institute of Geological Sciences, Edinburgh, and are available for the cost of reproduction. The data from the second line (WINCH) will be available in July 1983.

From 1 October 1982 D.H. Matthews will be a senior research associate in the University of Cambridge, Director of the BIRPS core group.

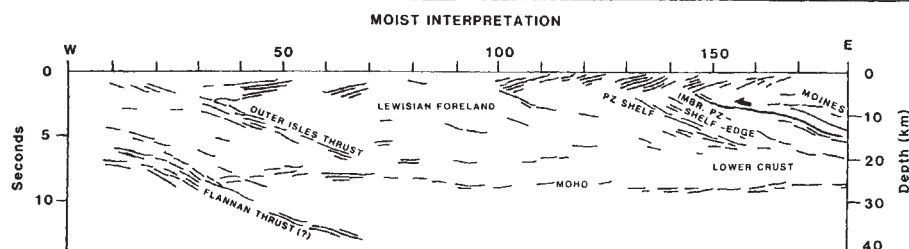


Fig. 2 Line drawing of reflection line-ups identified on MOIST (migrated data). The arrows indicate one plausible position of the Moine Thrust and the lettering in the border indicates one plausible geological interpretation². Migration was done for BIRPS by Shell UK.

almost to the Moho, but the Moine Thrust is less easy to identify. At the eastern end of the profile there are easterly dipping reflectors which appear to pass up into the boundary faults of rotated blocks of Devonian sediments and are interpreted as imbricate thrusts reactivated as normal faults during a later period of stretching. There is a close parallel here with easterly dipping reflectors seen at the seaward end of COCORP profiles across the Appalachians^{2,5}.

WINCH was intended to elucidate the Flannan Thrust and to cross the principal faults of the Caledonides. Processing of the data has begun, as has planning for profiling in financial year 1983–84 around Cornwall and probably across the English Channel towards Guernsey to investigate Hercynian structures, the Cornubian batholith and the Brioverian basement. It is intended to work with the French ECORS group which is forming to carry out deep seismic profiling in France.

A start has thus been made on investigating the deep geological structure under the British Isles using the only technique that promises to have sufficient resolution. Also, two short deep profiles have been shot by the deep-geology unit of the Institute of Geological Sciences on land and one by the University of Leicester. The oil industry has also begun to obtain some deep profiles at sea and on land. Resolution is limited by the low frequencies of the returns from deep reflectors: a frequency of 20 Hz in rock with sound velocity $V_p = 6 \text{ km s}^{-1}$ implies a wavelength of 300 m and the possibility, at a depth of 30 km, of resolving two reflecting points 75 km apart vertically and 2,100 m apart horizontally. The recording parameters are chosen to look at deep features so that the resolution of shallow events is less good than that of standard industry reflection profiling which is usually run for about 5 s two-way time. For resolving details of relatively shallow sedimentary strata a wealth of data obtained by the oil industry is held by the Department of Energy.

The main objectives of BIRPS are threefold. First, structural geology,

particularly to trace the attitude of major faults which are known at outcrop or subcrop to deep within the crust. Do the normal faults curve, to 'sole out' at relatively shallow depths? Are many thrust faults splays off sub-horizontal *decoulement* surfaces like the Brevard zone⁵ in the southern Appalachians? What is the shape of the Moho under sedimentary basins? Are all granites floored at about 6 s two-way time?

Second, continental rheology. A glance at a map of earthquake distribution is sufficient to convince that plate tectonics does not work in Eurasia: it is impossible to define aseismic plates delineated by lines of epicentres at their edges, as one can in the deep oceans, anywhere between Spain and Singapore. Is this because the hot continental lithosphere responds to stress by flowing, generating earthquakes only at shallow depths, or is it because the continents have such a long history of faulting that they can always respond to stress by reactivating some older fault surface, producing the widespread belt of

earthquakes with random mechanism solutions that we observe in Eurasia. In the latter case, we should commonly find faults extending right through the crust into the upper mantle. In this context, it is likely that depth or penetration could be extended deeper than the 15 s that we have so far tried. Attenuation is low in crystalline rocks ($Q \sim 400$) and signal-to-noise ratios are still workable at the bottom of the record if reflectors are there.

Third, imaging. There is a perceived need, both in industry and in academic deep seismics, for better techniques of imaging steeply dipping reflectors. Groups of theoretical seismologists are working on this problem both in industry and at Cambridge, Princeton and Stanford. We also need to compute the reflection profiles that we would expect from model structures, for comparison with observed profiles, if we are to understand how we get apparently flat-lying reflections, albeit discontinuous, from the foliation surfaces of the high-grade metamorphic rocks that we expect to occur deep in the continental crust and which we observe exposed at outcrop on the continental shields.

BIRPS will try to contribute to all three objectives, and perhaps in the future we may be able to profile abroad in areas of active mountain building. Working deep within the crust sounds esoteric, of interest only to academic structural geologists. But, like COCORP in the USA, to whom we are deeply indebted⁶, we are investigating the skeleton of the continent which supports and delimits the soft sediments of the oil basins. □



100 years ago

SCIENTIFIC EXPLORATION IN EGYPT

Now that we have embarked in a war in Egypt, it is to be hoped that steps will be taken to have a proper staff of scientific explorers attached to the army with facilities for conducting their investigations. There are periods of rest in a campaign during which soldiers and others may be usefully employed in conducting excavations at comparatively slight cost; and difficulties in the way of investigation, arising from the requirements of trade and industry, disappear in time of war. The deposits of the Delta require to be examined. The graves of the Nile Valley have to be connected with their animal remains. Much has to be done for the earliest and best period of Egyptian art, and the Stone Age of Egypt has to be fixed with certainty.

I trust also that we shall not rob Egypt of her antiquities to any great extent. It may be useful to complete our typical series to a limited extent, but if Boulak should be happily preserved, I hope it will be preserved for Egypt and not brought home. Nothing would serve

more to prove that we go there to civilise and not to rob. The means of communications are now so easy that all who are interested in Egyptology can see it there. Steam and railways have materially altered the requirements of education in this respect. Humanity, and British humanity in particular, now pours through all the great arteries of the world, and people observe and study more abroad than at home. The time has passed when antiquities should be regarded as trophies of war. It is no longer necessary for instruction to hoard up valuable specimens of foreign antiquities in European museums. So long as science has access to the materials of knowledge, that is all that it is necessary to bring away. Besides which, it should be remembered that the atmosphere of Egypt preserves antiquities in a way that no other climate can do; and when this fact hereafter becomes fully impressed upon the public mind, the time may come when subscriptions will be raised to take back obelisks and put them up again in their proper places; at any rate we have enough of them weathering and withering in smoke and damp. But the opportunity for exploration should not be lost. The French savants did their work thoroughly during their military operations in that country, and it would be shameful if, with the knowledge now at our disposal, the British expedition did not achieve more for the promotion of science than was effected by Napoleon half a century ago.

From *Nature* 26, 364, 17 August 1882.

1. Smythe, D.K. *et al.* *Nature* (in the press).
2. Brewer, J.A. & Smythe, D.K. (in preparation).
3. Bamford, D., Nunn, K., Prodehl, C. & Jacob, B. *Geophys. J. R. astr. Soc.* 54 (1), 43 (1978).
4. White, R.S., Jones, E.J.W., Hughes, V.J. & Matthews, D.H. *Tectonophysics* (in the press).
5. Cook, F.A., Brown, L.D. & Oliver, J.E. *Scient. Am.* 243 (4), 156 (1980).
6. Oliver, J.E. *Nature* 275, 485 (1978).

ARTICLES

A pregalactic Population III: primordial helium and heavy elements and the microwave background

P. W. Tarbet & M. Rowan-Robinson

Department of Applied Mathematics, Queen Mary College, Mile End Road, London E1 4NS, UK

The helium, heavy elements and dark remnants produced by a pregalactic generation of stars, Population III, are discussed quantitatively for three scenarios: (1) Population III is responsible for the distortion of the microwave background spectrum measured by Woody and Richards; (2) Population III objects generate the whole of the microwave background; (3) Population III provides the oldest metals found in our Galaxy today.

A NATURAL consequence of the isothermal density fluctuation picture is the possibility of a pregalactic generation of stars, Population III, where 'stars' includes objects up to 10^6 or 10^7 solar masses. Consequences of such a population have been considered elsewhere (see ref. 1 and refs therein). Here we investigate three different roles for Population III. First, to generate the distortion of the microwave background measured by Woody and Richards^{2,3}, as proposed in ref. 4. Second, to produce the entire microwave background, and the observed 'primordial' helium abundance. Third, to manufacture the first metals in galaxies.

These three possibilities are explored quantitatively with assumptions about the mass spectrum of the stars, the helium and heavy element yields of stars of different masses and the remnants.

Helium and metal yields of Population III

The mean mass density parameter of the Universe at the present epoch, $\Omega_0 (= 8\pi G\rho_0/3H_0^2 = (\rho_0/2 \times 10^{-29} \text{ g cm}^{-3}) (H_0/100)^{-2})$, can be considered to have three components; Ω_g , the contribution from visible matter, Ω_r , that due to dark matter, such as black holes and neutron stars, which we will assume are the remnants of Population III, and Ω_ν , from non-baryonic matter, such as neutrinos with mass $\Omega_0 = \Omega_g + \Omega_r + \Omega_\nu = \Omega_b + \Omega_\nu$, where Ω_b is the total baryonic contribution.

We assume that at some epoch in the early Universe a population of stars was born and subsequently died at a later epoch z_t (z_t is the redshift at which the objects cease to produce radiation, so in the case of black holes it is the epoch when they stop accreting at an appreciable rate).

If $Y(M)$ and $Z(M)$ are the fraction of a star's mass returned as helium and heavy elements respectively, then the fraction of visible matter in the form of helium and heavy elements is

$$Y = (1 - \Omega_{ej}/\Omega_g) Y_p + (8\pi G/3H_0^2 \Omega_g)(1 + z_t)^{-3} A \int_{M_1}^{M_2} Y(M) MN(M) dM$$

$$Z = (8\pi G/3H_0^2 \Omega_g)(1 + z_t)^{-3} A \int_{M_1}^{M_2} Z(M) MN(M) dM$$

where Y_p is the cosmic (pre-Population III) helium abundance, Ω_{ej} is the contribution to Ω_g that was cycled in pregalactic stars, H_0 is the Hubble constant and $AN(M) dM$ is the number-density of stars with masses in the range M to $M + dM$ at epoch z_t , having upper and lower limits of M_2 and M_1 respectively. We shall take $N(M) = M^{-\alpha}$. Similarly if $R(M)$ denotes the

fraction of a star's mass left behind as dark matter, then

$$\Omega_r = (8\pi G/3H_0^2) \cdot (1 + z_t)^{-3} \times A \left(\int_{M_1}^{M_b} R(M) MN(M) dM + (1 + \mu_B) \cdot \int_{M_b}^{M_2} R(M) MN(M) dM \right)$$

where it is assumed that the remnant is a black hole for $M > M_b$, and μ_B is the ratio of the mass accreted by a black hole to the mass of an initial black hole.

The coefficient A of the mass function is related to the radiation output of these objects. If they enhance the background by a factor β and $\varepsilon(M)$ is the fraction of a star's mass converted into radiation, then

$$\beta a T_0^4 (1 + z_t)^4 = A \int_{M_1}^{M_2} \varepsilon(M) MN(M) c^2 dM$$

$T_0(1 + z_t)$ being the temperature of the primordial cosmic background at epoch z_t . If the efficiency of producing radiation in the conversion of H to He is $\varepsilon_Y (\sim 0.007)$ and from He to metals is $\varepsilon_Z (\sim 0.003)$, then

$$\varepsilon(M) = \varepsilon_Y (Y(M) + Z(M) - Y_p) + \varepsilon_Z Z(M)$$

for the nucleosynthetic contribution; and for those stars which produce black holes at their deaths, which then accrete a fraction μ_B of their initial mass with efficiency ε_B

$$\varepsilon(M) = \varepsilon_B \mu_B R(M)$$

The value of $\varepsilon_B \mu_B$ is difficult to determine as the accretion rate depends on the temperature of the material being accreted

Table 1 The effect of varying the parameters M_2 , α , μ_B and H_0 on models producing the distortion of the microwave background

M_2	α	μ_B	H_0	Ω_g	Ω_b	$Y_{\min}$	Y_p	Z
10^6	1.0	5.0	50	0.043	0.100	0.259	0.256	2.85×10^{-5}
10^6	1.25	5.0	100	0.016	0.030	0.262	0.258	1.32×10^{-4}
10^5	0.75	5.0	50	0.375	0.432	0.275	0.268	1.63×10^{-5}
10^7	1.25	5.0	50	0.021	0.078	0.256	0.254	6.63×10^{-5}
10^6	1.0	1.0	50	0.147	0.243	0.268	0.263	4.16×10^{-5}
10^6	0.75	0.5	50	0.173	0.316	0.270	0.266	1.02×10^{-5}
10^6	0.75	0.1	50	0.769	1.294	0.282	0.277	1.15×10^{-5}

$Y_{\min}$ is the minimum net helium yield given by the parameters M_2 , α , μ_B , H_0 and Ω_g ; the corresponding Ω_b , Y_p and Z are also shown.

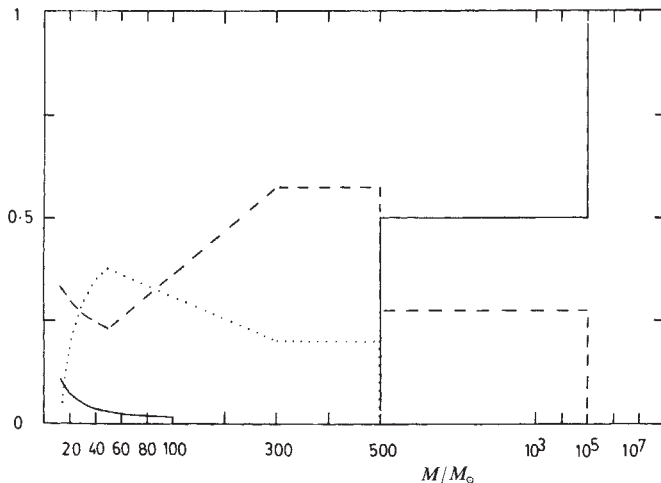


Fig. 1 The fraction of a star's mass ejected as helium ($Y(M)$, dashed curve) and heavy elements ($Z(M)$, dotted curve), and left as a remnant ($R(M)$, solid curve), as a function of the star's mass. Note that the mass scale changes at 100, 500 and $1,000M_{\odot}$.

which in turn is determined by the accretion, and the efficiency depends on the details of the flow and cooling processes⁵. The efficiency, ε_B , is of the order of 0.1 and if we consider $\varepsilon_{B\mu_B} = L_{ED}t/Mc^2$, where the Eddington luminosity $L_{ED} = 10^{38}(M/M_{\odot})$, and $t = 10^{17.5}\Omega^{-1/2}z_t^{-3/2}$, then $\varepsilon_{B\mu_B} = 0.02$ for $\Omega = 0.15$ and $z_t = 200$. But it was found that in the models having $z_t = 200$, $\varepsilon_{B\mu_B}$ could not be much less than 0.1, (see ref. 6), to avoid excessive amounts of matter or high yields from Population III.

Schramm and Steigman⁷ give $(0.9-3.5) \times 10^{-3}h_0^{-1}$, where $h_0 = H_0/100$, as a firm lower limit to the nucleon density parameter, based on observations of the light producing stars in the solar neighbourhood. This is taken as the lower limit for Ω_g .

The functions $N(M)$, $Y(M)$, $Z(M)$, $R(M)$

The function $N(M)$ is assumed to have the form $M^{-\alpha}$. For the solar neighbourhood $\alpha = 2.35$, the Salpeter value⁸, $M_1 = 0.01-0.1M_{\odot}$ and $M_2 = 60-100M_{\odot}$. In a zero metal abundance medium, cooling will be less efficient and fragmentation inhibited⁹. We may therefore expect M_1 to be increased (Silk⁹ argues that $M_1 \sim 20M_{\odot}$) and α reduced. For $\alpha < 2$, M_1 is not a very important parameter; $M_1 = 13M_{\odot}$ was adopted. The cocoon star mechanism that limits M_2 in the solar neighbourhood¹⁰ also does not apply in the metal-free primordial environment, and we assume that M_2 can extend up to the scale of the Jeans mass at decoupling.

In the range $12 < M/M_{\odot} < 100$ stars are believed to be able to burn all of their nuclear fuels through to iron without mishap¹¹. The core then undergoes hydrodynamic collapse but the implosion is reversed when nuclear densities are reached and a large amplitude bounce reflects the incoming kinetic energy, blowing off all layers external to the core in a supernova explosion^{12,13}. The exact details of core bounce are not yet known and the complex reaction networks of static and dynamic burning make yield determinations uncertain. Here, the helium yields of Arnett¹⁴ (scaled for an initial helium content of Y_p) are used, the formula of Weaver and Woosley¹⁵ is used for the metals, and the neutron star core remnants are taken to be $1.4M_{\odot}$.

Stars with initial mass greater than about $100M_{\odot}$ do not burn fuels all the way through to iron but encounter the e^+e^- pair instability and undergo core collapse¹⁶. The question is, can subsequent fast nuclear burning reverse this collapse or is the ultimate fate a black hole? Woosley and Weaver¹⁷ and Bond *et al.*¹⁸ calculate that stars having an initial mass much exceeding $300M_{\odot}$ (the uncertainty coming from the uncertainty in mass loss) will be unable to reverse this collapse, whilst those in the range $100-300M_{\odot}$ suffer complete disruption.

It has often been pointed out that massive stars ($M \geq 60M_{\odot}$)

are unstable to nuclear energized pulsations¹⁹⁻²¹. But studies by, for example, Appenzeller^{22,23}, Talbot^{24,25} and Papaloizou²⁶ imply that these pulsations do not completely disrupt the stars. For those of initial mass $>300M_{\odot}$ a substantial fraction of the star may be lost on a time scale much shorter than the hydrogen burning phase, whilst those below $100M_{\odot}$ suffer little from this instability²³.

Talbot and Arnett²⁷ looked at helium production from massive stars, with no initial helium, suffering mass loss due to nuclear pulsations and concluded that 17% of the star could be shed in the form of helium with a net helium yield of about 0.5 for stars $>300M_{\odot}$. However, other mechanisms (radiatively driven winds) may be important and the observations of Humphreys and Davidson²⁸ seem to point to the possibility that mass loss severely restricts the evolution of stars with initial masses $50-60M_{\odot}$ to cooler regions of the H-R diagram. It has been suggested²⁹ that OH-IRII sources with deep silicate absorption may be the remnants of massive stars which have suffered substantial mass loss.

Considering the uncertainties in mass loss and the upper limit of complete disruption the following simple approach was taken. Stars of mass $>M_a = 300M_{\odot}$ are assumed to lose half their mass due to winds and so (scaling for $Y_p = 0.25$) shed 0.275 of their mass as helium. The remnant then contributes to the yield after disruption to give a total fraction of 0.575 as helium and 0.2 as metals. Those above $M_b = 500M_{\odot}$ (up to $10^5M_{\odot}$) return the 0.275 wind component but the remnants collapse to black holes. For stars in the range $50-300M_{\odot}$, $Y(M)$ and $Z(M)$ were interpolated linearly between the values adopted at 50 and $300M_{\odot}$.

Objects exceeding $10^5M_{\odot}$ encounter a general relativistic instability during quasi-static collapse³⁰ and if they contain no heavy elements, nuclear burning cannot proceed rapidly enough to prevent catastrophic collapse³¹. Thus virtually the whole star collapses to a black hole on a short time scale.

The adopted forms for $Y(M)$, $Z(M)$ and $R(M)$ are summarized in Fig. 1.

Viable scenarios of Population III

(1) The distortion of the microwave background: To produce the spectrum observed by Woody and Richards², Negroponte *et al.*³² found it necessary to have z_t in the range 150-225, β in the range 0.2-0.4, and a metal abundance relative to the total mass of $2.5-7.5 \times 10^{-6}$ for $\Omega_0 = 1$ and $1-3 \times 10^{-5}$ for $\Omega_0 = 0.1$ ($H_0 = 100$), depending on the grain material. To test this type of model we take $z_t = 200$, $\beta = 0.25$ and look for values of the parameters Ω_g , H_0 , M_2 , α , μ_B , which give a heavy element abundance in the right range. Allowing for the uncertainties in the type of grain material and the observed distortion, we take this to be $\Omega_g Z = 5-50 \times 10^{-7}$ for $H_0 = 100$, and $10^{-6}-10^{-5}$ for $H_0 = 50$.

We have taken $H_0 = 50$ or 100; $\log_{10} M_2/M_{\odot} = 2(1/7)$; $\alpha = 0.5(0.25)2.5$; $\mu_B = 1, 5$; and in each case varied Ω_g from 0.001 to 1.

Table 2 Models in which Population III stars produce all the microwave background

a	M_2	α	μ_B	H_0	Ω_g	Ω_b	Y	Z
	10^5	0.5	5.0	50	0.091	0.394	0.223	9.26×10^{-5}
	10^5	0.5	5.0	100	0.023	0.099	0.223	9.26×10^{-5}
	10^6	0.75	1.0	50	0.013	0.522	0.220	3.51×10^{-4}
	10^6	0.75	1.0	100	0.003	0.130	0.220	3.51×10^{-4}
b	10^5	0.25	5.0	50	0.091	0.394	0.223	4.56×10^{-5}
	10^5	0.25	5.0	100	0.023	0.099	0.223	4.56×10^{-5}
	10^5	0.5	5.0	100	0.023	0.099	0.223	1.69×10^{-4}
	10^6	0.5	1.0	50	0.007	0.516	0.231	1.74×10^{-4}

a, The only models producing all of the microwave background and providing enough helium without overproducing heavy elements; b, as a but with $Z(M)$ increased at the expense of $Y(M)$ in the stellar mass range $300-500M_{\odot}$. Y is the net helium yield given by the parameters M_2 , α , μ_B , H_0 and Ω_g ; Z is the corresponding net heavy element yield.

Table 3 The effect of varying the parameters α , M_1 and M_2 on initial enrichment models

α	M_1	M_2	Ω_{III}	Ω_{ej}	Ω_{r}	Y_{III}	β
2.35	20	100	3.18×10^{-5}	3.06×10^{-5}	1.19×10^{-6}	8.37×10^{-5}	3.93×10^{-4}
1.0	20	100	2.96×10^{-5}	2.87×10^{-5}	8.33×10^{-7}	7.74×10^{-5}	3.91×10^{-4}
2.35	13	100	3.90×10^{-5}	3.70×10^{-5}	1.99×10^{-6}	1.08×10^{-4}	4.11×10^{-4}
2.35	30	100	2.88×10^{-5}	2.80×10^{-5}	8.13×10^{-7}	7.30×10^{-5}	3.85×10^{-4}
2.35	20	300	3.27×10^{-5}	3.18×10^{-5}	8.58×10^{-7}	9.95×10^{-5}	4.30×10^{-4}
2.35	20	500	3.38×10^{-5}	3.30×10^{-5}	8.04×10^{-7}	1.11×10^{-4}	4.54×10^{-4}

See text for definitions of the parameters and their functional dependence on Z , Ω_g and z_t .

Because the Jeans mass at decoupling is $\sim 10^6 M_\odot$ it is natural to look first for models with $M_2 = 10^6 M_\odot$ which produce the required $\Omega_g Z$. $M_2 = 10^6 M_\odot$, $H_0 = 50$, $\alpha = 1.0$ and $\mu_B = 5.0$ fits the required $\Omega_g Z$ and has a minimum helium yield of $Y = 0.259$ at $\Omega_g = 0.043$, $\Omega_b = 0.100$, $Z = 2.85 \times 10^{-5}$ and $Y_P = 0.256$. The dependence of Y (and Y_P) on Z and Ω_g for this model is shown in Fig. 2. For low values of Ω_g the net metal and helium yields from Population III are high but decrease as Ω_g increases. At higher Ω_g , and hence Ω_b , the primordial helium yield increases so Y starts to increase again. The model cited above is seen to have a net helium yield slightly higher than some recently deduced primordial values; $Y = 0.216 \pm 0.015$ (ref. 33), $Y = 0.228 \pm 0.014$ (ref. 34). As most of the helium comes from primordial nucleosynthesis this problem is shared by the standard model unless $\Omega_b < 0.03$ (ref. 7). However, these low values are disputed by Kunth³⁵, who deduces $Y = 0.243 \pm 0.010$.

We have used the work of Yang *et al.*³⁶ to determine Y_P , which in turn adopts the nucleosynthesis code of Wagoner³⁷. As discussed by Olive *et al.*³⁸, Y_P (and consequently Y) could be reduced by taking a lower value for the free neutron half life, $\tau_{1/2}$. Bondarenko *et al.*³⁹ have determined $\tau_{1/2} = 10.13 \pm 0.09$ min, compared with $\tau_{1/2} = 10.7$ used by Wagoner and $\tau_{1/2} = 10.56 \pm 0.09$ recommended in ref. 40. This could reduce Y_P by ~ 0.01 . If the τ -neutrino has a mass > 1 MeV/ c^2 , then the two-neutrino case rather than the three-neutrino case should be used in the helium nucleosynthesis calculation (see the dotted curve in Fig. 2). However, this case leads to implausibly high values of Ω_ν unless the mass of the τ -neutrino is greater than 2 GeV/ c^2 and this is inconsistent with the experimental upper limit of 250 MeV/ c^2 (ref. 41).

The deuterium in these models is given by

$$D = (1 - \Omega_{\text{ej}}/\Omega_g) D_P$$

and since, for the case cited, $\Omega_{\text{ej}} = 5.06 \times 10^{-4}$, $D = 0.988 D_P$, (almost identical to the standard model).

We consider the effect of varying the other parameters of our model. Changing H_0 to 100 decreases $\Omega_g Z$ by a factor of 4 whilst the required range in $\Omega_g Z$ is reduced by a factor of 2. The net effect of changing H_0 is therefore small. To recover a model giving $\Omega_g Z$ in the required range we need to increase α from 1 to 1.25 when H_0 is changed from 50 to 100. The dependence of $\Omega_g Z$ on α is more significant. Increasing α to 1.25 with other parameters unchanged still leaves $\Omega_g Z$ in the required range, but $\alpha = 1.5$ gives $\Omega_g Z = 5.70 \times 10^{-5}$ and $\alpha = 0.5$ gives $\Omega_g Z = 2.55 \times 10^{-8}$. Since the radiation is mainly from accreting black holes, the upper end of the mass spectrum, the effect of reducing M_2 is to require a lower value of α to maintain $\Omega_g Z$ in the required range. The yields are also increased as M_2 is reduced because the contribution to the radiation from objects returning processed material is higher. It was found that M_2 must exceed $10^4 M_\odot$ unless $\alpha < 0.5$. Lowering μ_B increases the yields from Population III and such models also require more matter (higher Ω_r to provide the required radiation), hence Y_P is increased. Table 1 illustrates the effects of varying the parameters M_2 , α , μ_B and H_0 .

Thus a pregalactic stellar population can provide a distortion of the microwave background as in the scheme outlined in ref. 4, but to give acceptable values of Y and Z the mass function must be flat and extend to super-massive objects, so that most of the light is generated by accreting black holes.

(2) Generation of the whole of the microwave background:

As pointed out by Carr⁶, if a pregalactic generation of stars could have generated 25% of the microwave background, could it not be possible for them to have generated all of it? In this case $\beta = 1.0$ and we take $T_0 = 2.9$ K so as to give a total energy density equal to that observed by Woody and Richards^{2,3}. The primordial helium yield now has to be generated entirely by Population III stars (nucleosynthesis needs to be suppressed in a cold Universe by taking the lepton-to-baryon ratio > 1.3 to prevent most of the Universe being synthesized into helium and heavy elements)⁴². If we are still trying to fit the claimed distorted spectrum then the required $\Omega_g Z$ is as in the previous case, and we take $z_t = 200$. We have also looked at $z_t = 1,000$ because, in the Carr model, the extra 75% of the background radiation produced in these models has to be generated at this earlier epoch for thermalization by free-free processes to be effective. (The Rowan-Robinson *et al.*⁴ scheme requires only that the radiation be produced by $z_t \sim 200$. The constraints on $\Omega_g Z$ are due to the optical depth in dust back to $z_t \sim 200$ required in this scheme, so are unchanged.)

The only models giving the required $\Omega_g Z$ and enough helium are given in Table 2a. All other models (with $\mu_B = 1.0$ or 5.0) satisfying the constraints on $\Omega_g Z$ and Z , from old Population II, produced too little helium.

Models with lower μ_B were found. With $M_2 = 10^6 M_\odot$, $\alpha = 0.75$ and $\mu_B = 0.5$, we found $Y = 0.220$ and $Z = 3.51 \times 10^{-4}$ with $\Omega_g = 0.007 h_0^{-2}$, $\Omega_b = 0.197 h_0^{-2}$ ($\Omega_r = 0.190 h_0^{-2}$), but the only two cases with acceptable yields for $\mu_B = 0.1$ were $M_2 = 10^6 M_\odot$, $\alpha = 0.5$ and $M_2 = 10^7 M_\odot$, $\alpha = 0.75$ which had $\Omega_r = 0.697 h_0^{-2}$ and $\Omega_r = 0.670 h_0^{-2}$, respectively, which are rather high. We also found acceptable models with $z_t = 1,000$ but these also required high densities ($\Omega_r > 0.6 h_0^{-2}$) or had $< 1\%$ of the total baryonic matter in visible form today, so are rather implausible.

Note that none of these models produce any deuterium and it is therefore necessary to appeal to some *ad hoc* astrophysical mechanism for this (see ref. 43). However, as there are models that imply that the Carr scenario can simultaneously produce all of the microwave background, explain the Woody and Richards distortion with the mechanism proposed in ref. 4, and give the correct helium yield and Population II metals, it would perhaps be worthwhile to investigate the predicted background spectrum of these models in more detail, taking into account the staggered way in which the radiation from different stellar masses is generated. The capacity of free-free processes to thermalize the bulk of the radiation has been disputed by Wright⁴⁴. He argues that for adequate thermalization at long wavelengths, needle-shaped conducting grains are required, with a length-to-width ratio of at least 210; a similar mechanism is proposed in ref. 45.

(3) Initial enrichment: The need for a non-zero initial metal content for the Galaxy has been argued by many authors (see refs 46–48) and most recently by Bond⁴⁹ who deduces a mean metal abundance of 0.0181 $Z_\odot$ for Population II and an initial metal content of $2.51 \times 10^{-3} Z_\odot$.

It has been seen that models (1) and (2) can provide this initial enrichment, but here we suppose the radiation requirement on Population III to be minimal and try to produce the required metal abundance without adding significantly to the primordial helium.

Individual galaxies are believed to condense out at about

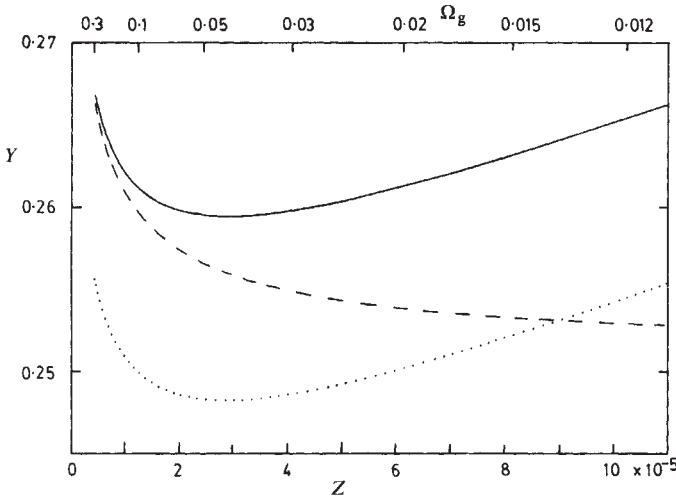


Fig. 2 The total helium abundance, Y (solid curve) and the primordial helium abundance, Y_p (broken curve), plotted as a function of the heavy element abundance, Z (lower scale). The contribution to Ω_b due to visible matter, Ω_g , is labelled on the upper scale. The dotted curve shows the two-neutrino case. Model parameters are given in the text.

redshifts 10, so z_i is relaxed to this value. However, β is not a relevant parameter in these models as we are not concerned with the effect of the stars on the microwave background. Instead we consider the fraction of the Universe that needs to go into a pregalactic stellar generation to provide the initial metals. We define Ω_{III} , the density parameter today for these objects if they had not ejected any of their mass (or accreted any).

$$\Omega_{III} = (8\pi G/3H_0^2) \cdot (1+z_i)^{-3} A \int_{M_1}^{M_2} MN(M) dM$$

First let us consider $M_1 = 20M_\odot$ (ref. 9), $\alpha = 2.35$ (the Salpeter⁸ value), and $M_2 = 100M_\odot$. For this set of parameters we find $\Omega_{III} = 3.18 \times 10^{-5}$ ($Z/10^{-4}$) ($\Omega_g/0.1$), $\Omega_{ej} = 3.06 \times 10^{-5}$ ($Z/10^{-4}$) ($\Omega_g/0.1$), $\Omega_t = 1.19 \times 10^{-6}$ ($Z/10^{-4}$) ($\Omega_g/0.1$), $Y_{III} = 8.37 \times 10^{-5}$ ($Z/10^{-4}$), (where Y_{III} is the net helium yield from Population III) and $\beta = 3.93 \times 10^{-4} h_0^2$ ($Z/10^{-4}$) ($\Omega_g/0.1$) ($11/1+z_i$).

It is therefore possible to provide the initial metal content of galaxies without resorting to a flat mass function extending to SMOs, and with only a very small amount of the Universe being cycled in pregalactic stars. Such models have a negligible effect on primordial helium and deuterium abundances and on the microwave background.

The effect on these values of variation of the parameters α , M_1 and M_2 is illustrated in Table 3.

Only for high values of M_2 and low values of α , comparable with those required in cases (1) and (2), does β , and hence the distortion of the microwave background, become significant.

Discussion

Because of uncertainties in Fig. 1 we now discuss the effects on our results of varying some of the values adopted for the yield functions.

(1) We have redone the calculations for scenarios (1) and (2) ($\mu_B = 1.0$ and 5.0 , $z_i = 200$) varying the limit M_a , where a star

loses half its mass due to pulsations, between 300 and $500M_\odot$, and the limit M_b , where black hole formation occurs, between 300 and $1,000M_\odot$. Varying these limits was found to have no qualitative effect on our previous conclusions and a negligible effect on the quantitative results.

(2) Klapp⁵⁰ has performed numerical calculations of the evolution of very massive stars and obtained yields which differ from those we have adopted (Fig. 1). We therefore consider the effects of changing $Y(M)$ and $Z(M)$ in two different ways: (i) increasing $Z(M)$ from 0.2 to 0.45 in the mass range 300 – $500M_\odot$ at the expense of $Y(M)$, so more helium is cycled through to metals than assumed previously (we still linearly interpolate $Y(M)$ and $Z(M)$ between the values adopted at 50 and $300M_\odot$), and (ii) increasing $Z(M)$ from 0.0 to 0.02 in the mass range 500 – $10^5M_\odot$, so a small amount of the material processed into metals is able to diffuse up to the surface of the star and is shed before the collapse to a black hole occurs (see ref. 50).

(i) The effect of increasing $Z(M)$ at the expense of $Y(M)$ over a small section of the mass spectrum is, generally, that α has to be reduced fractionally for a given model to keep $\Omega_g Z$ in the required range. However, in our case (1) above (producing the distortion of the microwave background) it was found that our best model ($M_2 = 10^6M_\odot$, $H_0 = 50$, $\alpha = 1.0$ and $\mu_B = 5.0$) is still viable and gave $Y = 0.259$ at $\Omega_g = 0.041$, $\Omega_b = 0.098$, $Z = 4.99 \times 10^{-5}$ and $Y_p = 0.256$. For models producing all of the microwave background (case (2)) the only cases giving $\Omega_g Z$ in the required range and enough helium, without violating the constraints on Z are given in Table 2b.

(ii) In this case $Z(M)$ has been increased over a large range of the mass spectrum and it is therefore necessary to reduce α more substantially than in case (i) to recover viable models. Our best model now has $\Omega_g Z = 2.13 \times 10^{-5}$ and is therefore invalid. Reducing α from 1.0 to 0.5 , with all other parameters held constant, brings $\Omega_g Z$ back into the required range and gives $Y = 0.256$ at $\Omega_g = 0.021$, $\Omega_b = 0.078$, $Z = 2.94 \times 10^{-4}$ and $Y_p = 0.254$. The effect of varying the parameters of this model are again as previously discussed, but it was found that M_2 must exceed $10^5M_\odot$ unless $\alpha < 0.25$. No models were found that produced all of the microwave background and enough helium without overproducing metals. These models are much more dependent on the helium to metal production ratio of Population III than those producing only the distortion of the microwave background, because for the latter most of the helium is provided by nucleosynthesis in the fireball phase of the Universe.

We conclude that small changes in our adopted yield functions (case (1)) only affect the details of our results, not the general conclusions. But if Klapp's conclusion that even very massive stars which collapse to black holes return some of their mass in the form of heavy elements is correct, then models trying to produce all of the microwave background and the initial helium from Population III are untenable.

There are several viable and interesting roles for Population III and the best hope of testing these ideas observationally is improved measurements of the spectrum of the microwave background near the peak.

P.W.T. is supported by an SERC research studentship. We thank B. Carr, S. Woosley and J. Klapp for providing results before publication.

- Rowan-Robinson, M. & Tarbet, P. in *Progress in Cosmology* (ed. Wolfendale, A.) (Reidel, Dordrecht, 1982).
- Woody, D. P. & Richards, P. L. *Phys. Rev. Lett.* **42**, 925–929 (1979).
- Woody, D. P. & Richards, P. L. *Astrophys. J.* **248**, 18–37 (1981).
- Rowan-Robinson, M., Negroponte, J. & Silk, J. *Nature* **281**, 635–638 (1979).
- Carr, B. J. *Mon. Not. R. astr. Soc.* **189**, 123–136 (1979).
- Carr, B. J. *Mon. Not. R. astr. Soc.* **195**, 639–668 (1981).
- Schramm, D. N. & Steigman, G. *Astrophys. J.* **243**, 1–7 (1981).
- Salpeter, E. E. *Astrophys. J.* **121**, 161–167 (1955).
- Silk, J. *Astrophys. J.* **211**, 638–648 (1977).
- Kahn, F. D. *Astr. Astrophys.* **37**, 149–162 (1974).
- Weaver, T. A., Zimmerman, G. B. & Woosley, S. E. *Astrophys. J.* **225**, 1021–1029 (1978).
- Arnett, W. D. *Ann. N.Y. Acad. Sci.* **336**, 366–379 (1980).
- Van Riper, K. A. *Astrophys. J.* **232**, 558–571 (1979).
- Arnett, W. D. *Astrophys. J.* **219**, 1008–1016 (1978).
- Weaver, T. A. & Woosley, S. E. *Ann. N.Y. Acad. Sci.* **336**, 335–357 (1980).

- Barkat, Z., Rakavy, G. & Sack, N. *Phys. Rev. Lett.* **18**, 379–381 (1967).
- Woosley, S. E. & Weaver, T. A. *Supernovae* (eds Rees, M. & Stoneham, R. J.) (Reidel, Dordrecht, 1982).
- Bond, J. R., Arnett, W. D. & Carr, B. J. *Supernovae* (eds Rees, M. & Stoneham, R. J.) (Reidel, Dordrecht, 1982).
- Stothers, R. in *Stellar Evolution* (eds Hong-Yee Chiu & Muriel, A.) (Massachusetts Institute of Technology Press, 1972).
- Schwarzschild, M. & Harm, R. *Astrophys. J.* **129**, 637–646 (1959).
- Stothers, R. & Simon, N. R. *Astrophys. J.* **160**, 1019–1029 (1970).
- Appenzeller, I. *Astr. Astrophys.* **5**, 355–371 (1970).
- Appenzeller, I. *Astr. Astrophys.* **9**, 216–220 (1970).
- Talbot, R. J. *Astrophys. J.* **163**, 17–27 (1971).
- Talbot, R. J. *Astrophys. J.* **165**, 121–138 (1971).
- Papalopoulos, J. C. B. *Mon. Not. R. astr. Soc.* **162**, 169–187 (1973).
- Talbot, R. J. & Arnett, W. D. *Nature Phys. Sci.* **229**, 150–151 (1971).
- Humphreys, R. M. & Davidson, K. *Astrophys. J.* **232**, 409–420 (1979).

29. Rowan-Robinson, M. *Mon. Not. R. astr. Soc.* (in the press).
30. Zeldovich, Ya. B. & Novikov, I. D. *Relativistic Astrophysics* Vol. 1 (University of Chicago Press, 1971).
31. Fricke, K. J. *Astrophys. J.* **183**, 941–958 (1973).
32. Negroponte, J., Rowan-Robinson, M. & Silk, J. *Astrophys. J.* **248**, 38–46 (1981).
33. French, H. B. *Bull. Lick Obs.* No. 863 (1979).
34. Lequeux, J., Peimbert, M., Rayo, J. F., Serrano, A. & Torres-Peimbert, S. *Astr. Astrophys.* **80**, 155–166 (1979).
35. Kunth, D. in *Cosmology and Particles* (eds Audouze, J. et al.) 241 (Frontières, Paris 1981).
36. Yang, J., Schramm, D. N., Steigman, G. & Rood, R. T. *Astrophys. J.* **227**, 697–704 (1979).
37. Wagoner, R. V. *Astrophys. J.* **179**, 343–360 (1973).
38. Olive, K. A., Schramm, D. N., Steigman, G., Turner, M. S. & Yang, J. *Astrophys. J.* **246**, 557–568 (1981).
39. Bondarenko, L. N., Kurguzov, V. V., Prokof'ev, Yu. A., Rogov, E. V. & Spivak, P. E. *JEPT Lett.* **28**, 303–306 (1978).
40. Wilkinson, D. H. in *Proc. of the Erice School on Nuclear Astrophysics* (1980).
41. Cowsik, R. in *Cosmology and Particles* (ed. Adouze et al.) 157 (Frontières, Paris, 1981).
42. Carr, B. J. *Astr. Astrophys.* **60**, 13–26 (1977).
43. Epstein, R. I. *Astrophys. J.* **212**, 595–601 (1977).
44. Wright, E. L. *Astrophys. J.* **255**, 401–407 (1982).
45. Rana, N. C. *Mon. Not. R. astr. Soc.* **197**, 1125–1138 (1981).
46. Adouze, J. & Tinsley, B. M., *Ann. Rev. Astr. Astrophys.* **14**, 43–79 (1976).
47. Truran, J. W. & Cameron, A. G. W. *Astr. Space Sci.* **14**, 179–222 (1971).
48. Hartquist, T. W. & Cameron, A. G. W. *Astr. Space Sci.* **48**, 145–158 (1977).
49. Bond, H. E. *Astrophys. J.* **248**, 606–611 (1981).
50. Klapp, J. thesis, Univ. Oxford (1982).

Vitrification of pure liquid water by high pressure jet freezing

Erwin Mayer & Peter Brüggeller

Institut für Anorganische und Analytische Chemie, Universität Innsbruck, A 6020 Innsbruck, Austria

The vitrification of pure liquid water by projecting a thin jet of liquid water at high speed into a liquid cryomedium is reported. The influence of the experimental parameters on the cooling rate and the devitrification of the jet-frozen vitrified material have been investigated. A structural difference between vitrified liquid water and amorphous solid water prepared from the vapour phase is inferred from a comparison of the X-ray diffraction patterns.

FINDING a method for vitrifying pure liquid water is important for several reasons. The investigation of vitrified liquid water will help to elucidate the structure of liquid water because the effects of thermal and static disorder can be largely separated^{1–4}. It might help in interpreting the anomalous properties of super-cooled water^{5–8}. Vitrification of biological systems without cryoprotectants is another application that could enable studies to be made of morphology without freezing artefacts and of the mechanism of freezing damage^{9,10}. When applied to dilute aqueous solutions complete vitrification means the simultaneous matrix isolation of dissolved species. This extends the well known matrix isolation from the gas phase¹¹ to nonvolatile solutes such as ionic systems or biomolecules and allows, among others, spectroscopic investigations at low temperatures in the dissolved state. Attempts to form vitreous ice by rapid cooling of liquid water have invariably led to formation of ice I_h (ref. 12). (Pryde and Jones¹³ did report a heat capacity change of rapidly cooled water at 126 K which they attributed to a glass transition, but could not reproduce this result in subsequent experiments.) We have recently described the complete vitrification of liquid water and dilute aqueous solutions by jet-freezing of aqueous droplets distributed in *n*-heptane as an emulsion, where the combination of emulsification and high cooling rate was essential for vitrification¹⁴. We have since increased the cooling rate by orders of magnitude and report here the vitrification of pure liquid water without emulsification by injecting a thin fluid jet of liquid water at high pressures *in vacuo* into a liquid cryomedium. This is an extension of the jet-freezing method developed by Bachmann and Schmitt for freeze etching¹⁵. Our technique differs from that of splat-cooling on a cold metal surface by using a liquid cryomedium, and gives for liquid water and dilute aqueous solutions much higher cooling rates¹⁴. [It has been reported that D. R. Uhlmann has vitrified liquid water by splat-quenching (ref. 16 in our ref. 6). However, to our knowledge no report has appeared since then.]

Conditions for vitrification

We have investigated the following parameters, considered to be important for a high cooling rate, with the ESR method described previously¹⁴: aperture diameter from 5 to 100 μm , working pressures between 10 and 400 atm, effect of the liquid cryomedium and of stirring. Small aperture diameter together with high working pressure were found to be most important for high cooling rates. With increasing working pressures the

liquid jet is broken up by air resistance closer and closer to the exit hole of the aperture and dispersed into a fine mist. Therefore vacuum is essential for working at pressures above 100 atm. Liquid ethane, propane and butene-1 were investigated as cryomedia and gave for small apertures very similar cooling rates. Figure 1a shows the X-ray diffraction pattern of a jet-frozen sample of liquid water using the best conditions for vitrification emerging from hundreds of experiments: jetting through a 10- μm aperture with 400 atm working pressure (the

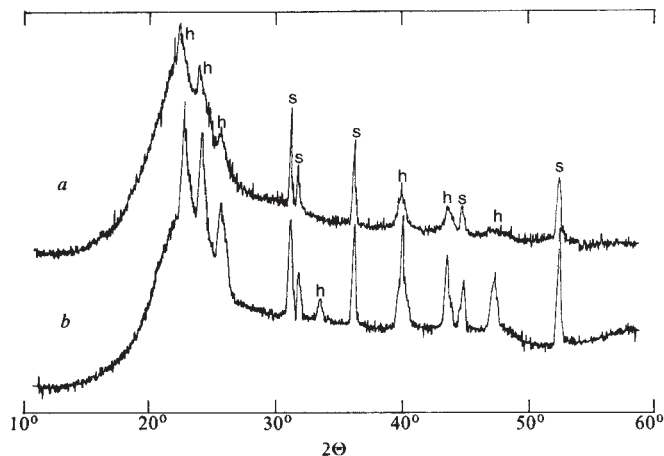


Fig. 1 X-ray diffraction patterns (CuK α , 90 K) of jet-frozen pure liquid water with: a, 400 atm; b, 100 atm working pressure (s, substrate reflexes; h, reflexes from crystalline ice). Samples were jet-frozen as follows: between 2 and 5 ml of pure liquid water were jetted through an aperture with a 10- μm opening (Balzers Union no. 16001), using working pressures of at most 400 atm, into 120 ml of vigorously stirred liquid propane cooled to 80 K. Liquid consumption at 100 atm: 0.4 ml min⁻¹, at 400 atm: 1.1 ml min⁻¹. The working pressures were generated with a hydraulic pump (type Maximator MSF 72), plugging of the 10- μm aperture was prevented by two filters (a Nupro 2- μm metal filter and a Millipore high pressure filter unit with a 0.2 μm filter). The experiment was started at normal pressure and vacuum applied as quickly as possible. Starting and stopping the jet already *in vacuo* gave identical X-ray diffraction patterns, but seemed to give apertures that clogged more easily. The suspension was filtered at liquid nitrogen temperature and the remaining mixture of frozen water and liquid cryomedium was transferred quickly with a liquid-nitrogen cooled spoon into the precooled apparatus. All operations at normal pressure were carried out under a cover of dry nitrogen gas. At liquid-nitrogen temperatures special care has to be taken to prevent accumulation of liquid oxygen in the cryomedium. The X-ray data were recorded on a Kristalloflex 4 (Siemens), using a vertical flat sample holder and samples between 0.5 and 1 mm thick.

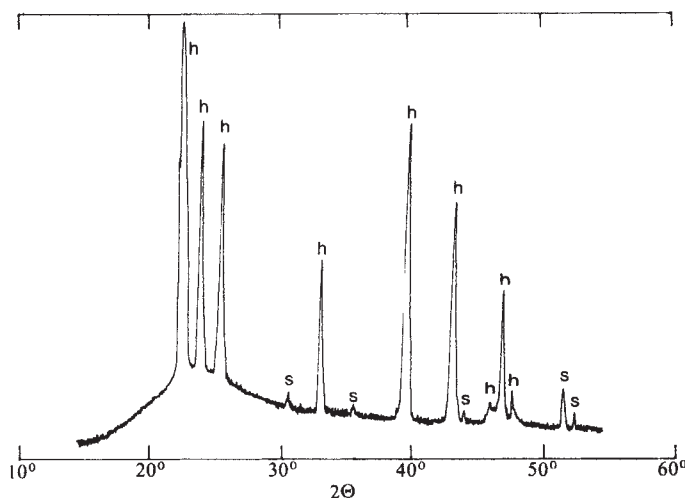


Fig. 2 X-ray diffraction pattern ($\text{CuK}\alpha$, 90 K) of jet-frozen pure liquid water with 10 atm working pressure, showing the contribution of the liquid cryomedium at $2\theta = 23^\circ$ (s, substrate reflexes; h, reflexes from ice I_h). The experimental conditions and the instrument gain were the same as conditions for Fig. 1 except for the working pressure and an aperture with a 20- μm opening. A 20- μm aperture was used because we could not start the jet through a 10- μm aperture with only 10 atm working pressure. However, from jet-freezing experiments with 50 atm we know that apertures with 10 and 20 μm openings give comparable cooling rates in this range of working pressures. With 10 atm working pressure the comparatively slow jet is not broken up by air over a long distance and thus the cooling rate is not increased by jetting *in vacuo*. The sharp reflexes are due to ice I_h , the intensity changes at small angles being caused by texture.

highest pressure used) into vigorously stirred propane at 80 K *in vacuo*. The X-ray diffraction pattern contains only a small amount of crystalline—mainly hexagonal—ice and is dominated by the broad feature characteristic of vitreous or microcrystalline substances. The effect of high working pressures is clearly seen by comparison with Fig. 1b where the pressure was reduced to 100 atm, using otherwise jet-freezing conditions identical with the experiment of Fig. 1a, and where the amount of hexagonal ice increases strongly. In several experiments we obtained nearly identical X-ray diffraction patterns with the same set of jet-freezing conditions which means that this method of vitrification is highly reproducible.

The vitrified liquid water inevitably contains some liquid cryomedium which cannot be removed without partial devitrification. The effect of the liquid cryomedium on the X-ray diffraction pattern of vitrified liquid water is demonstrated by Fig. 2, where the pattern of jet-frozen water using 10 atm working pressure only is shown with approximately the same intensity as in Fig. 1a, b: by differential thermal analysis (DTA) and by the ESR method reported previously¹⁴ we confirmed that jet-freezing with 10 atm working pressure gives only crystalline ice. The X-ray diffraction pattern contains a broad peak centred at $2\theta = 23^\circ$ in addition to the strong reflexes of ice I_h and is very similar to the pattern of a mixture of the liquid cryomedium with crystalline ice. The broad peak in Fig. 2 is therefore caused by the liquid cryomedium¹⁶ and not by vitrified liquid water. The comparatively strong X-ray intensity of the liquid hydrocarbon might at first seem surprising. The small absorption coefficient for X rays of the solid water and the resulting transparency of the sample provide a simple explanation. In addition a matrix effect of the liquid cryomedium leading to strongly enhanced Compton scattering¹⁷ and a particle size effect¹⁸ might have a role. In any case the jet-freezing conditions for Fig. 2 are useful as a 'null-experiment', showing clearly in comparison with Fig. 1 the effect of high working pressures on the relative X-ray intensities of crystalline and amorphous material. We do not consider it useful to correct the X-ray diffraction pattern of vitrified liquid water for the effect of the cryomedium quantitatively because the amount of liquid and thus the size of the broad peak in Fig. 2 varied

somewhat, and the possible additional effects considered above cannot be evaluated.

The vitrified material is definitely liquid water and not water vapour because: (1) the liquid enters the cryomedium as a continuous fluid jet of $\sim 10\ \mu\text{m}$ diameter. Typically between 2 and 5 ml of liquid water were jetted using a vacuum of 1 torr total pressure. The jetting time depends on the working pressure and the diameter of the aperture and was varied between 1 and 5 min. Therefore the amount of water vapour present during the experiment is negligible in comparison to the total amount of liquid water jetted. (2) We had tried to vitrify liquid water in the form of a very fine mist of small water droplets using working pressures of 1 to 2 atm only. The water droplets were generated by an ultrasonically vibrating 10- μm orifice similar to the aerosol generator of Berglund and Liu¹⁹. These experiments were carried out with a very small flow rate over considerable periods of time and therefore represent conditions where the amount of water entering the liquid cryomedium from the vapour phase becomes increasingly important. However, we have found no indication that material was vitrified using this arrangement. (3) The ESR spectra of jet-frozen 0.10 M aqueous CuCl_2 solutions depend strongly on the working pressures and were used as a good indicator for relative cooling rates¹⁴. We found that with a 200-atm working pressure the jet-frozen ESR spectrum is already nearly identical to the ESR spectrum of a perfectly vitrified 0.10 M CuCl_2 solution containing 50% glycerol for vitrification. While this does not prove that the jet-frozen solid has been vitrified completely, it still proves that the solute CuCl_2 has been matrix isolated and hence, that the liquid has experienced very high cooling rates.

Our jet-freezing conditions give a clear improvement, in terms of the amount of the vitrified material, over the previously reported vitrification of emulsified liquid water droplets¹⁴ where we had obtained yields of ~ 13 and 16% for pure water and a 0.10 M CuCl_2 solution by DTA. These yields had been only barely noticeable in the X-ray diffraction pattern as a broadening underneath the sharp reflexes of the crystalline material between $2\theta = 20^\circ$ and 28° (see Fig. 1a in ref. 14), whereas in Fig. 1a we have only a small amount of crystalline material on top of the broad feature. The determination of the relative amount of vitrified liquid water by DTA does not have any meaning for samples prepared as described in Fig. 1a, because the DTA method is not accurate enough, and the crystalline material comes mainly from condensation of water vapour during the preparation and transfer of the sample to the X-ray camera.

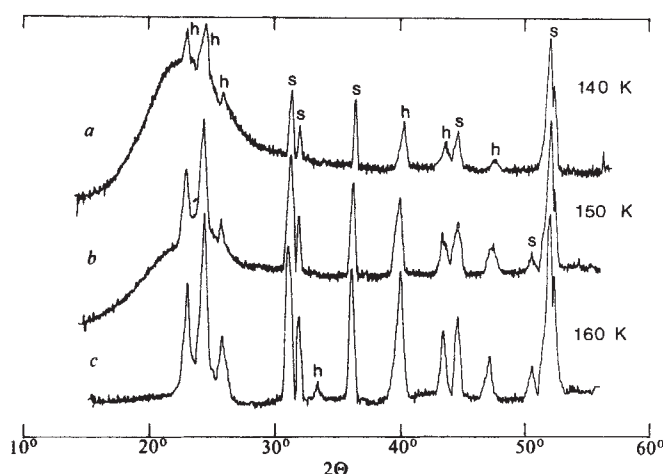


Fig. 3 X-ray diffraction patterns ($\text{CuK}\alpha$) of jet-frozen pure liquid water warmed consecutively for 30 min to a, 140 K; b, 150 K; c, 160 K. The sample was prepared as described in Fig. 1 with 400 atm working pressure except that liquid butene-1 was used as cryomedium at 100 K instead of propane. From 90 to 140 K the relative intensities of amorphous to crystalline reflexes were approximately constant, only the overall intensity decreased, probably due to movement of the viscous paste on the vertical sample holder with increasing temperature.

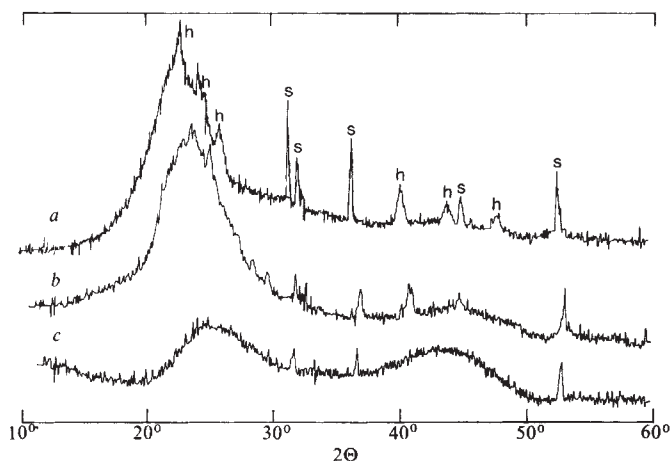


Fig. 4 X-ray diffraction patterns ($\text{CuK}\alpha$, 90 K) of: *a*, jet-frozen pure liquid water prepared as described for Fig. 1*a* and thus containing some liquid propane; *b*, $(\text{H}_2\text{O})_{\text{as}}$ milled with liquid propane; *c*, $(\text{H}_2\text{O})_{\text{as}}$ without propane. $(\text{H}_2\text{O})_{\text{as}}$ was prepared as described in ref. 22 by vapour condensation at 77 K on a Cu-plate from a bath for the reservoir of water held at 210 K, using high vacuum conditions. For comparison with the pattern in Fig. 4*a* $(\text{H}_2\text{O})_{\text{as}}$ was mixed thoroughly at 80 K with 20 ml of liquid propane, filtered and the remaining paste transferred to the precooled apparatus, using the same technique as for jet-frozen liquid water samples. As the pattern of $(\text{H}_2\text{O})_{\text{as}}$ in Fig. 4*c* contains no reflexes of crystalline impurities, the amount of crystalline material in Fig. 4*b* represents the contamination of vitrified samples by crystalline ice due to condensation of water vapour during the experimental procedure.

We have investigated the devitrification and found that within the time scale of our X-ray measurements (30 min at every consecutive 10 K, starting at 90 K) the formation of ice I_c starts at 150 K (Fig. 3). This agrees very well with the devitrification of a vitrified 0.10 M CuCl_2 solution investigated previously within a similar time scale using ESR where a devitrification temperature of 153 K was deduced from signal broadening¹⁴. We found that the cryomedium used for vitrification affects the devitrification temperature. The sample giving the X-ray diffraction patterns in Fig. 3 had been prepared by jetting liquid water into butene-1 and therefore contained some butene-1. Vitrified liquid water containing some propane had started to devitrify already at 130 K. We think that this difference comes from the higher vapour pressure of liquid propane (1 torr at 144 K) and that devitrification is possibly induced by evaporating propane. At low temperatures the X-ray diffraction patterns of vitrified samples prepared with butene-1 or propane as cryomedia, using otherwise identical jet-freezing conditions, are very similar. The X-ray diffraction patterns of jet-frozen samples stored at 77 K for 3 months did not show any change.

Comparison with $(\text{H}_2\text{O})_{\text{as}}$

In Fig. 4 the X-ray diffraction pattern of vitrified liquid water containing some liquid propane (Fig. 4*a*) is compared with the pattern of amorphous solid water prepared by condensation from the vapour phase (called $(\text{H}_2\text{O})_{\text{as}}$ in accordance with Rice¹, Fig. 4*b*, *c*). Figure 4*c* shows the published X-ray diffraction pattern of $(\text{H}_2\text{O})_{\text{as}}$ (refs 3, 20). Figure 4*b* the pattern of $(\text{H}_2\text{O})_{\text{as}}$ milled with liquid propane. A comparison has meaning only between the patterns of samples treated both with liquid cryomedium and has to be made between the patterns shown in Fig. 4*a*, *b*. A difference might possibly be found at high angles where the broad feature in the pattern of $(\text{H}_2\text{O})_{\text{as}}$ centred around $2\theta = 44^\circ$ appears to be less intense or absent in the pattern of vitrified liquid water. For further comparison we have tried to remove the cryomedium by washing the sample with liquid methane. However, this procedure leads to partial devitrification of the sample. In Fig. 5 the X-ray diffraction pattern of vitrified liquid water washed with liquid methane (Fig. 5*b*) is compared with the pattern of $(\text{H}_2\text{O})_{\text{as}}$ treated identically (Fig. 5*a*), showing again for vitrified liquid water no clear amorphous peak at $2\theta = 44^\circ$ underneath the sharp reflexes of

crystalline ice in contrast to the pattern of $(\text{H}_2\text{O})_{\text{as}}$. The X-ray diffraction pattern of the devitrified sample in Fig. 5*c* can be used as background of Fig. 5*b*. However, only a presentation of the X-ray data in the form of the structurally sensitive part of the scattered X-ray intensity will make a clear differentiation possible, and we cannot do this because we have not been able to compensate for the effect of the liquid cryomedium.

Discussion

The influence of high working pressures on the amount of vitrified material and thus on the cooling rate is certainly connected with the speed of the liquid jet. For 10, 100 and 400 atm working pressures approximate jet speeds of 20, 100 and 250 m s^{-1} can be calculated from the diameter of the 10- μm aperture and consumption of liquid. The effect of jet speed can be easily demonstrated by jetting at room temperature into paraffin oil with a viscosity similar to liquid propane at 80 K. With 10 atm working pressure the liquid jet hardly penetrates the surface, whereas with 100 atm working pressure penetration goes up to 20–30 mm. Therefore the liquid jet passes within a given time with increasing jet speed a larger volume of cryomedium, causing more efficient heat transfer from water to the cryomedium and giving higher cooling rates. In addition, vigorous stirring of the cryomedium enforces a circular movement onto the liquid jet, probably breaking up the jet into small droplets.

It is impossible to measure the cooling rates obtainable with high pressure jet-freezing. Fletcher¹² estimated the cooling rate necessary for vitrifying a micrometre-sized droplet of water to be $>10^{10} \text{ K s}^{-1}$ from the theory of the homogeneous nucleation of freezing, whereas Uhlmann²¹ calculated a cooling rate of $>10^7 \text{ K s}^{-1}$. We do not know how the liquid jet breaks up in the cryomedium at high pressures, but we consider it likely that the jet breaks up into droplets of approximately the size of the jet diameter. Therefore a minimum cooling rate of 10^{10} K s^{-1} is obtained from Fletcher's value and of 10^7 K s^{-1} from Uhlmann's. As the minimum cooling rate depends strongly on droplet diameter and was calculated for 1- μm droplets^{12,21} much higher cooling rates would result for larger droplet diameters. For a 10- μm droplet minimum cooling rates of 10^{13} and 10^{10} K s^{-1} follow from Fletcher's and Uhlmann's calculations. An idea about the distribution of cooling rates can be obtained from an inspection of the frozen products. We had

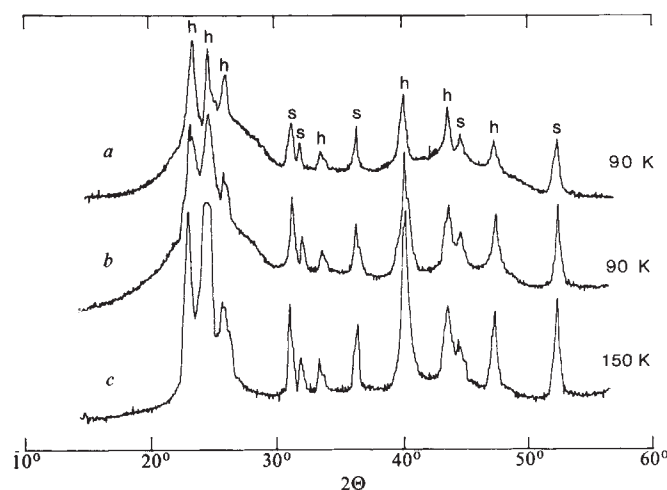


Fig. 5 X-ray diffraction patterns ($\text{CuK}\alpha$) of: *a*, $(\text{H}_2\text{O})_{\text{as}}$ from Fig. 4*b* washed with liquid methane; *b*, jet-frozen liquid water from Fig. 4*a* washed with liquid methane; *c*, the sample in *b* devitrified by warming up to 150 K. To remove liquid propane or butene-1 the samples were washed with liquid methane, filtered at $\sim 90 \text{ K}$ as described in Fig. 1 for propane, and the remaining methane pumped off at 77 K from the sample. This procedure leads inevitably to strong contamination and/or partial devitrification of the samples as shown by a comparison of Fig. 5*a* with Fig. 4*b* and *c*. The reflex at $2\theta = 24.3^\circ$ in Fig. 5*c* was truncated.

expected that with increasing working pressure, increasing amounts of ice I_c and then of vitrified liquid water would be formed. This was frequently observed for the preparation of $(H_2O)_{as}$ from the vapour phase^{20,22}. However, Fig. 1b shows that the crystalline fraction consists mainly of ice I_h in addition to the vitrified material. Thus the liquid jet apparently experiences either low cooling rates, giving ice I_h , or very high cooling rates leading to vitrified water. We believe that this is due to the vortex of the vigorously stirred cryomedium.

The thermodynamic conditions necessary for vitrifying liquid water have been discussed recently by Angell and Tucker⁸. The argument is based on the incompatibility of an extrapolated value of the glass transition temperature of vitrified liquid water with Kauzmann's entropy paradox²³ which states that a liquid can be supercooled only by a limited number of degrees before it loses all its excess entropy over the crystal, that is before its total entropy becomes equal to that of the crystal. This presents a problem, especially for liquid water, because in supercooled water a drastic increase of C_p occurs with decreasing temperature⁵ among other anomalies. To resolve the entropy paradox Angell and co-workers postulated the existence of a λ -type heat capacity anomaly at 228 K associated either with some liquid state mechanical instability or with the cooperative formation of a long-range hydrogen-bonded network with decreasing temperature⁶⁻⁸. Angell and Tucker speculate that liquid water departs during cooling (assuming a rate of 10^8 K s⁻¹) from the metastable heat capacity already below the maximum observed value (which has been recorded at 235 K (refs 6, 24)) due to the 'freezing out' of the 'anomalous fluctuations' (see Fig. 5 in ref. 8). From the C_p versus temperature curve we wonder whether liquid water does not diverge at much higher temperatures—thereby avoiding the anomalous C_p increase

altogether—and lose its internal equilibrium than that suggested by Angell and Tucker, considering the extreme cooling rates discussed above. In analogy to SiO_2 the vitrified solid is supposed to have different densities depending on the temperature at which internal equilibrium was established before quenching^{25,26}. However, a determination of the density of vitrified liquid water is not possible at present due to the presence of the liquid cryomedium.

The similarity of structure between vitrified liquid water and $(H_2O)_{as}$ has been disputed recently. Rice^{1,3,27} and Angell⁸ believe that there is a continuity of state between liquid water and $(H_2O)_{as}$, that $(H_2O)_{as}$ is essentially liquid water, especially because the observed glass transition temperature of $(H_2O)_{as}$ is identical with the extrapolated value for vitrified liquid water. This hypothesis is the basis of the random network model of water recently proposed by Rice *et al.*^{4,28}. In contrast Johari²⁹ argues on thermodynamic grounds mentioned above that a discontinuity of states exists, and that therefore "the structure of $(H_2O)_{as}$ is different from the one towards which supercooled water would tend". We have tried without success to observe a glass transition temperature in our vitrified water samples. A structural difference between vitrified liquid water and $(H_2O)_{as}$ could possibly emerge from a comparison of the X-ray diffraction patterns (see Figs 4 and 5). We do not know yet if the absence of the broad peak at $2\theta = 44^\circ$ in vitrified liquid water is real or is only an artefact caused by the presence of the liquid cryomedium. According to Narten and Levy³⁰ a change of the double maximum in the structure functions of liquid water around $2.5 \text{ \AA}^{-1}$ into a single peak corresponds to a disappearance of the maxima and minima between 3.5 and 8 Å in the correlation functions.

We thank Professor E. Schnell for help with the X-ray work.

Received 10 December 1981; accepted 10 May 1982.

1. Olander, D. S. & Rice, S. A. *Proc. natn. Acad. Sci. U.S.A.* **69**, 98–100 (1972).

2. Rice, S. A. & Sceats, M. G. *J. phys. Chem.* **85**, 1108–1119 (1981).

3. Narten, A. H., Venkatesh, C. G. & Rice, S. A. *J. chem. Phys.* **64**, 1106–1121 (1976).

4. Rice, S. A. *Topics curr. Chem.* **60**, 109–200 (1975).

5. Angell, C. A. in *Treatise on Water* Vol. 7 (ed. Franks F.) (in the press).

6. Angell, C. A., Shuppert, J. & Tucker, J. C. *J. phys. Chem.* **77**, 3092–3099 (1973).

7. Speedy, R. J. & Angell, C. A. *J. chem. Phys.* **65**, 851–858 (1976).

8. Angell, C. A. & Tucker, J. C. *J. phys. Chem.* **84**, 268–272 (1980).

9. Mazur, P. *Science* **168**, 939–949 (1970).

10. Wolstenholme, G. E. W. & O'Connor, M. *The Frozen Cell* (Churchill, London, 1970).

11. Becker, E. D. & Pimentel, G. C. *J. chem. Phys.* **25**, 224–228 (1956).

12. Fletcher, N. H. *Rep. Prog. Phys.* **34**, 913–994 (1971).

13. Pryde, J. A. & Jones, G. O. *Nature* **170**, 685–688 (1952).

14. Brüggeller, P. & Mayer, E. *Nature* **288**, 569–571 (1980).

15. Bachmann, L. & Schmitt, W. W. *Proc. natn. Acad. Sci. U.S.A.* **68**, 2149–2152 (1971); in *Freeze Etching* (eds Benedetti, E. L. & Favard, P.) 73–79 (Soc. Frac. Microsc. Electron, Paris 1973).

16. Stewart, G. W. & Mannheimer, M. Z. *anorg. allg. Chem.* **171**, 61–72 (1928).

17. Sahores, J. J. *Adv. X-ray Analysis* **16**, 186–197 (1973).

18. Berry, P. F., Furuta, T. & Rhodes, J. R. *Adv. X-ray Analysis* **12**, 612–632 (1969).

19. Berglund, R. N. & Lui, B. Y. H. *Envir. Sci. Technol.* **7**, 147–153 (1973).

20. Dowell, L. G. & Rinfret, A. P. *Nature* **188**, 1144–1148 (1960).

21. Uhlmann, D. R. *J. Non-Cryst. Solids* **7**, 337–348 (1972).

22. Hobbs, P. *Ice Physics* (Clarendon, Oxford, 1974).

23. Kauzmann, W. *Chem. Rev.* **43**, 219–256 (1948).

24. Rasmussen, D. H., MacKenzie, A. P., Tucker, J. C. & Angell, C. A. *Science* **181**, 342–344 (1973).

25. Kanno, H. & Angell, C. A. *J. chem. Phys.* **73**, 1940–1947 (1980).

26. Bruckner, R. J. *Non-Cryst. Solids* **5**, 281 (1971).

27. Rice, S. A., Bergren, M. S. & Swingle, L. *Chem. Phys. Lett.* **59**, 14–16 (1978).

28. Rice, S. A. & Sceats, M. G. *J. phys. Chem.* **85**, 1108–1119 (1981).

29. Johari, G. P. *Phil. Mag.* **35**, 1077–1090 (1977).

30. Narten, A. H. & Levy H. A. *J. chem. Phys.* **55**, 2263–2269 (1971).

The molecular basis of DNA–protein recognition inferred from the structure of cro repressor

D. H. Ohlendorf*, W. F. Anderson†, R. G. Fisher*, Y. Takeda‡ & B. W. Matthews*

*Institute of Molecular Biology and Department of Physics, University of Oregon, Eugene, Oregon 97403, USA

†MRC Group on Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7

‡Chemistry Department, University of Maryland, Baltimore County, Catonsville, Maryland 21228, USA

Recognition by cro repressor protein of its specific DNA binding sites appears to occur via multidentate hydrogen bonds between amino acid side chains of the protein and base-pair atoms in the major groove of right-handed B-form DNA. Most of the sequence-specific interactions between cro and DNA, as well as a number of sequence-independent ones, are mediated by a two- α -helical unit which appears to be common to many proteins that regulate gene expression.

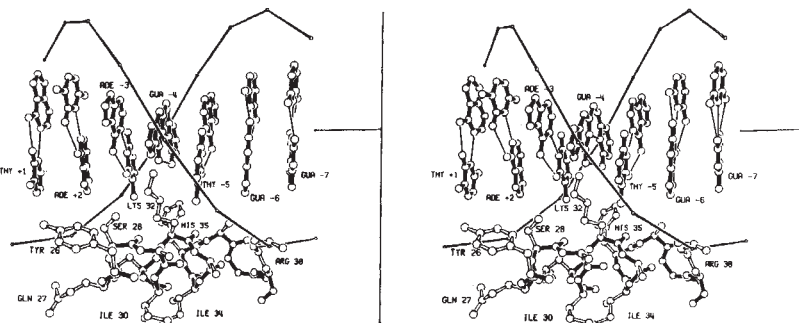
THE structures of three proteins that specifically bind to DNA have recently been determined, namely, the cro repressor protein (hereafter called 'cro') from bacteriophage λ ¹, the catabolite gene activator protein (CAP) from *Escherichia coli*² and the amino-terminal fragment of the cI repressor protein (cI) from bacteriophage λ ³. It has been proposed that these proteins bind to DNA with protruding α -helices penetrating^{1,2} or partly penetrating³ the major grooves of the DNA (in the

case of CAP, the DNA has been postulated to be left-handed²).

In this article we propose a detailed model for the complex between cro and DNA. The model is consistent with the relative binding affinities of cro for the six operator sites within bacteriophage λ at which it binds specifically, and can also be used to rationalize known changes in binding affinity caused by mutations within these sites.

Cro is one of a number of proteins that participate in the

Fig. 1 Stereo drawing illustrating the initial alignment of residues 26–38 of cro repressor within the major groove of the DNA. Only the left half of the 17-base-pair binding site is shown. The vertical line at the right of the figure is the common protein and DNA twofold-symmetry axis which relates the other half of the binding site to that shown here. The horizontal line at the right coincides with the long axis of the DNA. The conformations of the individual protein side chains are those determined from the crystal structure of cro repressor protein. The base sequence is that for O_R3, with numbering as defined in the text. For clarity the details of the DNA backbone have been omitted and replaced by lines connecting the phosphates.



commitment to either the lytic or lysogenic pathway during the development of bacteriophage λ (for review see refs 4–6). Cro recognizes six binding sites on the bacteriophage genome, three within the left operator region (O_L1, O_L2 and O_L3) and three within the right operator region (O_R1, O_R2 and O_R3). Each binding site consists of 17 base pairs with approximate twofold symmetry. The base sequences of all six sites are similar but not identical⁷.

Starting parameters

Coordinates for the DNA were derived from the fibre diffraction analysis of B-DNA by Arnott and colleagues (refs 8, 9 and S. Arnott, personal communication). Initially, the rise per base-pair, z , was set at 3.4 Å and the twist per base-pair, θ , was set at 34.6° (10.4 base-pairs per turn). These parameters are representative of values obtained by others^{10–12}.

The coordinates for cro were based on the previously reported structure determination¹. The quality of the electron density map has been improved by a process of electron density averaging and phase extension¹³ in which the final agreement between observed and electron-density-averaged structure factors was 22.3% at 2.8 Å resolution (R.G.F., D.H.O. and B.W.M., unpublished results). Also the cro structure has been partially refined, the coordinates used here being taken from the present stage of refinement, for which the crystallographic residual is 27% at 2.2 Å resolution (23% at 2.8 Å resolution) (D.H.O. and B.W.M., unpublished). The estimated error in the protein coordinates is about 0.4 Å.

The protein dimer was placed so that its twofold axis of symmetry coincided with that of the DNA¹, leaving two adjustable parameters, one the separation, d , between the protein and the DNA, and the other the rotation, τ , of the protein dimer relative to the DNA about the common protein-DNA symmetry axis. In addition to the above four parameters, z , θ , d and τ , adjustments were allowed in the radius of curvature, r , of the DNA. Because of the twofold symmetry of the protein-DNA complex, bending was constrained to occur only about points on the dyad axis.

Model building

Our general approach was to carry out model building on an MMS-X computer graphics facility¹⁴ and to supplement this with energy minimization using the program EREF of Jack and Levitt¹⁵. Use of computer graphics permitted the screening of a large number of potential interactions between the protein and the DNA while energy minimization ensured that the stereochemistry of a proposed DNA-protein complex was reasonable.

The energy minimization procedure gives the calculated energy for the system under consideration, and, as such, provides an indication of the relative strength of a postulated cro:DNA complex. However, the energy minimization algorithm does not take account of solvent, and has other limitations which prevent its use for the reliable estimation of absolute interaction energies. Therefore, in attempting to optimize a given cro:DNA complex, we used energy minimization as a guide, but supplemented this with additional criteria such as the adherence of postulated hydrogen bonds to acceptable stereochemistry.

There has been considerable speculation that the free energy of stabilization of DNA:protein complexes comes primarily from relatively nonspecific interactions¹⁶ while specificity derives largely from hydrogen bonds between protein side chains and the parts of the base pairs exposed in the grooves of the DNA^{16–19}. Therefore, in model building we attempted to maximize the number of hydrogen bonds and other interactions between cro and the DNA, and assumed that each 'mispositioned' (that is, unsatisfied) hydrogen bond decreased the affinity of the DNA-protein complex, although it is understood that other factors such as shape complementarity, hydrophobic and dipolar interactions and the role of solvent are also important. Since the binding constant, K , for the cro:DNA interaction is related to the free energy of the interaction, ΔG , by the formula $K = \exp(-\Delta G/RT)$, it follows that each improperly positioned hydrogen bond donor or acceptor, which reduces the free energy by about 1 kcal mol⁻¹ (refs 16, 20), results in a reduction of the binding constant by a factor of approximately seven.

The presumed complexes between cro and DNA were initially subjected to energy minimization in which the protein backbone was kept rigid and only those side chains which appeared to interact with the DNA were allowed to move. A specialized version of EREF was used which allowed the explicit specification of hydrogen bonds. The five interaction parameters z , θ , d , τ and r (see above) were then systematically varied in small increments to determine which values gave the best value for the calculated energy of interaction. (This stepwise approach was necessary because EREF tends to make changes in local geometry, but not in the global interaction parameters.) In the final stages of the energy minimization, the local geometry of the DNA was allowed to vary, and, in addition, small adjustments were permitted in the protein backbone.

Interaction of cro with DNA

The initial structure determination of cro suggested that it bound to the DNA with a pair of dyad-related helices penetrating the major groove of the DNA¹. This overall complementarity between the surface of the protein and the DNA not only suggests the approximate alignment of the protein relative to the DNA, but (short of making major conformational changes in the protein which we shall consider below), makes it sterically difficult to devise alternative interaction schemes which seem plausible.

Since both cro and the DNA binding site have approximate twofold symmetry it is, in the first instance, sufficient to consider the interaction between one monomer of cro and half of the binding site. We have numbered the base pairs within each half-site from 1 to 9 starting at the distal base pair. Bases +1 and -1 constitute a base pair, and so on. The '+' base is the one for which the inward direction of numbering is from 5' to 3'.

The initial alignment of cro (Fig. 1) suggested that the amino acid side chains most likely to participate in interactions within the major groove of the DNA were Tyr 26, Gln 27, Ser 28, Asn 31, Lys 32, His 35 and Arg 38. Based on the framework of the potential hydrogen bonds and the cro:DNA geometry suggested by the energy calculations, it is possible to construct a set of rules which can be used to predict the hydrogen bonds

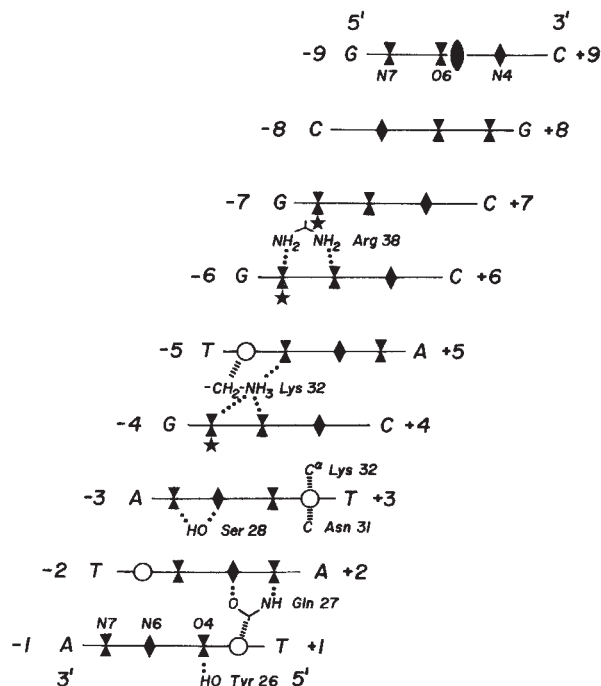


Fig. 2 Schematic representation, following Woodbury *et al.*²⁰, of the presumed sequence-specific interactions between cro and the parts of the base pairs exposed within the major groove of the DNA. The direction of view is imagined to be directly into the major groove of the DNA with the base-pairs seen edge-on. The dyad symbol within the topmost base pair indicates the center of the overall 17-base-pair binding region. The symbols are as follows: ∇ , hydrogen bond acceptor; $\blacklozenge$, hydrogen bond donor; $\bigcirc$, methyl group of thymine; $\star$, guanine N7 which is protected from methylation when cro is bound^{22,23}. Presumed hydrogen bonds between cro side chains and the bases are indicated (...). Apparent van der Waals contacts between cro and the thymine methyl groups are shown (|||||). Other van der Waals contacts are not shown.

expected when cro interacts with a given sequence of DNA. The proposed rules for hydrogen bonding are as follows:

(1) Arg 38 forms two hydrogen bonds to O6 and N7 of guanine -6. If base -6 is an adenine then Arg 38 makes a hydrogen bond with either the N7 of this adenine or the N7 of a purine at -7. If base -6 is a thymine then Arg 38 can make no hydrogen bonds to this base because of the blocking methyl group on C6 but can form one hydrogen bond to N7 of a purine at -7 (that is, a net loss of one hydrogen bond).

(2) Lys 32 N ϵ forms one hydrogen bond with either a purine N7 or an O4 of thymine at position -5. In addition, it forms a second hydrogen bond with O6 of guanine -4 and possibly a third hydrogen bond with N7 of guanine -4.

(3) Ser 28 O γ forms two hydrogen bonds with N6 and N7 of adenine -3 or, alternatively, one hydrogen bond to N7 of guanine, N4 of cytosine, or O4 of thymine at -3. In the case of a thymine at position -3 there is a second hydrogen bond from the N6 of the adenine at +3 to the carbonyl oxygen of the serine.

(4) Gln 27 forms two hydrogen bonds to N6 and N7 of an adenine at +2. In all other cases it forms a single hydrogen bond.

(5) Tyr 26 O η forms a hydrogen bond with either O4 of thymine +1 or N7 of a purine at +1, but cannot be placed in a stereochemically acceptable position to accept a hydrogen bond from N6 of adenine or N4 of cytosine.

As shown in Table 1, this model for the interaction between cro and DNA is consistent with the known relative affinities of cro for its six binding sites and for mutant sites as well.

The presumed sequence-specific hydrogen-bonding between cro and O_R3 is shown in Fig. 2 using a schematic representation based on that proposed by Woodbury *et al.*²⁰ and is shown in stereo in Fig. 3. The apparent interactions between the protein and the DNA backbone, which are not sequence-specific, and, presumably, provide much of the overall energy of interaction, are summarized in Fig. 4. The overall conformation of cro repressor when bound to its specific binding site is illustrated in Fig. 5.

In the proposed cro:O_R3 complex, there are about 20 sequence-specific hydrogen bonds (Fig. 2) as well as many sequence-independent interactions (Fig. 4). These interactions could easily account for the observed dissociation constant of about 10⁻¹⁰ M (refs 21, 22).

Because cro is known to bind most tightly to O_R3 (refs 23-25) we used this for the initial model building and for most of the energy calculations. After refinement of the cro:O_R3 complex, the DNA had a mean z of 3.48 $\pm$ 0.09 Å with a range of 3.30 Å to 3.56 Å and a mean twist θ of 34.7 $\pm$ 3.2° with a range of 26.9° to 39.8°. These values compare well with previously reported values for these parameters^{10,26,27}.

The optimization procedure suggests that the DNA is bent to a radius of about 75 Å. This degree of curvature is not excessive, as might at first be supposed. In the crystallized B-DNA dodecamer CGCGAATTCGCG, the DNA bends to a radius of 110 Å (ref. 28). Also, when associated with proteins in the nucleosome, DNA adopts a radius of about 45 Å (ref. 29). It can be estimated that the energy required to bend a 17 base-pair segment of DNA to a radius of 75 Å is between 0.85 and 1.7 kcal mol⁻¹ (refs 30, 31), an amount of energy approximately equivalent to the formation of one hydrogen bond per monomer. As an alternative to the bending of the DNA, an essentially equivalent result could be obtained by a 'hinge-bending' motion of the two cro monomers in the vicinity of the antiparallel β -sheet strands, Glu 54-Lys 56. These residues form part of the C-terminal 'arms' (residues 52-60) which appear to be the major determinant in stabilizing the cro dimer. Other than the contacts mediated by these two arms there are only very tenuous interactions between the respective monomers, suggesting that a 'hinge-bending' motion could easily occur (Fig. 5; see also Figs 3, 4 and 6 of ref. 1).

Specific and nonspecific cro:DNA interactions

The proposed model for the complex between cro and DNA suggests that interactions occur predominantly within two regions of the protein, namely residues 15-38, that is, essentially the α_2 and α_3 helices, and residues 54-66 which include part of the third β -sheet strand and the carboxyl terminus of the molecule.

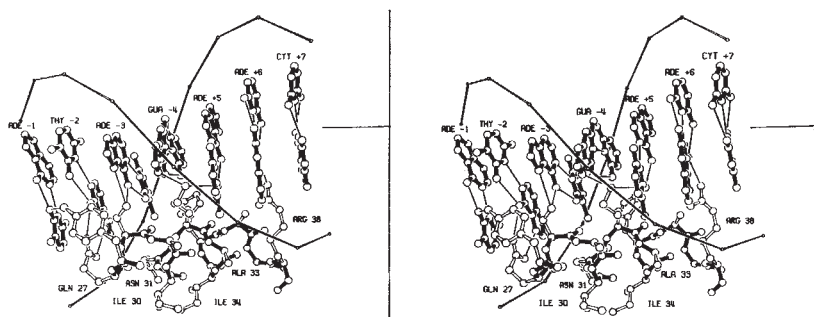


Fig. 3 Stereo drawing (Compare Fig. 1) showing the presumed hydrogen bonding of cro to base pairs in the major groove of O_R3, following model building and energy refinement.

Table 1 Cro binding to different DNA sites

Site	Sequence	Mutant	Affinity relative to O _R 3	Predicted H-bonds relative to O _R 3
O _R 3	¹ TATC ⁵ ACCG ⁹ CAAG ¹³ GGATA ¹⁷ ATAGTGGCGTTCCCTAT G C T A T A A T G C C G		1	0
		v3C	1/10	-1
		r1	1/3	0
		r2	<1	0
		r3	<1	0
		c12		-1
		c10	1/10	-1
O _R 2	TAACACCGTGCGTGTTG ATTGTGGCACGCACAAC G C A T AA TT T A		1/8	-1
		v3C	1/80	-2
		vi or vH		-2
		virC23	1/400	-3
		vN or vC34	<1/8	-3
O _R 1	TACCTCTGGCGGTGATA ATGGAGACCGCCACTAT G C A T A T C G T A		1/8	-1
		vR18		-2
		vs326	1/40	-2
		v3		-1
		vs387	<1/8	-2
		vC1	1/40	-2
O _L 1	TACCACTGGCGGTGATA ATGGTGACCGCCACTAT G C T A A T		~1/2	-1
		v101		-2
		v2		-2
		v003		-2
O _L 2	TATCTCTGGCGGTGTTG ATAGAGACCGCCACAAC A T		~1/2	-1
		v305		-1
O _L 3	AACCATCTGCGGTGATA TTGGTAGACGCCACTAT ¹⁷ ¹ ⁵ ⁹ ¹³		~1/10	-2
Amino acid	26 Lys Arg -28 32 38	Arg Lys 26 38 32 -28		

DNA sequences and relative affinities are shown for the six 17-base pair sites to which cro binds specifically in phage λ . Mutations in these sites with known relative affinities for cro binding are also shown. (Data taken from refs 23, 24.) The relative binding affinities are to be compared with the change in the number of sequence-specific hydrogen bonds between cro and the DNA site as predicted from the model for cro binding proposed in the text. If the model is correct, and if hydrogen bonding is a major determinant in the specificity of cro: DNA recognition, then the loss of successive hydrogen bonds ought to correspond to successive reductions in the affinity of binding by about a factor of seven.

The α_3 helix (residues 27–36) penetrates the major groove of the DNA and appears to contribute most of the sequence-specific interactions (Figs 2, 3). Interactions of Lys 32 and Arg 38 with bases in positions -4, -5 and -6 account for the observed protection from methylation of guanine N7 at these positions^{4,23,24}. In the presumed complex between cro and O_R3 (Figs 2, 3) there is no hydrogen bonding at the N7 of guanine -7. This is, perhaps, surprising in view of the observed protection of this nitrogen from methylation. However, the close

proximity of the guanidinium group of Arg 38 to guanine -7 (Fig. 3) might sterically inhibit the access of dimethyl sulphate to this base.

The α_2 helix (residues 15–23) appears to make a number of sequence-independent interactions with the outermost part of the DNA binding site (Fig. 4). Altogether, residues in the α_2 and α_3 helices may contribute as many as five hydrogen bonds to the phosphate backbone. In the model, there are no sequence-specific interactions between cro and the central three

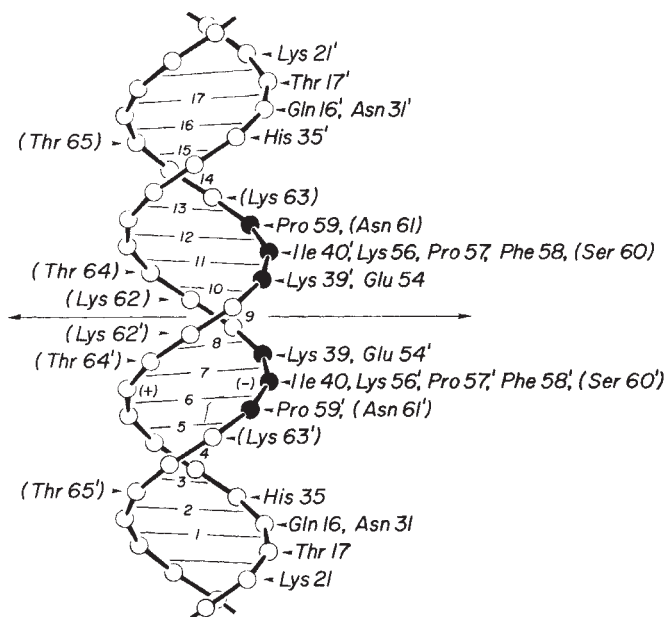


Fig. 4 Schematic representation of the contacts and interactions which are presumed to occur between cro and the phosphates in the DNA backbone. In the presumed cro:DNA complex, at least one close approach occurs between a given phosphate group and the named amino acids. Those phosphates for which ethylation interferes with cro binding²³ are drawn solid. Numbering of the base pairs and the signs (+) and (-) correspond to the identification used in the text. Amino acids belonging to 'lower' and 'upper' cro monomers related by the horizontal dyad axis are non-primed and primed, respectively. Amino acids in parentheses are those at the C-terminus of the molecule which are disordered in the crystal (residues 62–66) or for which the backbone conformation has been adjusted (residues 60, 61) to improve the contact with the DNA.

base pairs of the binding site, consistent with the observation that, for these three base-pairs, nitrogens N3 of the adenines exposed in the minor groove and N7 of the guanines exposed in the major groove are susceptible to methylation by dimethyl sulphate^{4,23,24}.

Residues Glu 54–Lys 56 of the third β -strand, together with the twofold-related strand in the adjacent cro monomer, are located close to the centre of the binding site and lie parallel to the minor groove, but, in part because of the bending of the DNA, do not lie within the minor groove (compare refs 1, 32). Rather, their contact with the DNA appears to be limited to interactions between the side chains and the phosphate backbone (Fig. 4). In the model for the presumed cro:DNA complex, the phosphate at position -6 (between base pairs 6 and 7 in Fig. 4) and possibly one or two solvent atoms or cations are surrounded by protein. The two neighbouring phosphates also appear to be partly 'covered' by protein (see Fig. 4), consistent with the observation that ethylation of these three phosphates interferes with the formation of the cro:DNA complex^{4,23}. In the presumed cro:DNA complex, it appears that the seven C-terminal residues, including Lys 62 and Lys 63, could lie along the minor groove of the DNA with the side chains interacting with the phosphate backbone (Figs 4, 5). Since the five C-terminal residues are disordered in the crystals of the native protein¹ it has to be emphasized that the details of such interactions remain very speculative.

Inferences for protein–DNA recognition

The proposed complex between cro and DNA has a number of features which may be of general significance for recognition by proteins of specific sites on DNA.

The initial determination of the cro structure¹ revealed a pair of twofold-related α -helices arranged in such a way that they could be placed within successive major grooves of right-handed Watson–Crick B-form DNA. This result gave credence to pre-

vious suggestions that α -helices of proteins might bind within the major groove of B-DNA^{33–36}.

The present study strongly suggests that sequence-specific recognition of DNA is mediated by hydrogen bonds between protein side chains and the parts of the base pairs exposed within the grooves of the DNA as has been advocated by a number of authors^{16,19,35–38}. In particular, it is of interest to note that the specificity of recognition appears to be enhanced by multiple interactions between a given side chain and appropriate hydrogen bond donors and acceptors on the base pairs. The apparent bidentate interaction of Gln 27 with nitrogens N6 and N7 of adenine +2 as well as the interaction of Arg 38 with O6 and N7 of guanine -6 (Figs 2, 3) are as anticipated^{18,39}. Ser 28 and Lys 32 are also presumed to participate in multiple interactions. Such interactions had not been postulated previously, although von Hippel and colleagues^{16,19} have suggested that bridging interactions could occur between atoms on successive base pairs, as seems to be the case with cro, in some instances. Also Drew and Dickerson⁴⁰ have observed a solvent molecule (molecule 66 in their numbering scheme) which forms hydrogen bonds with both the N6 and N7 nitrogens of an adenine, precisely as is proposed to occur between O^γ of Ser 28 and adenine -3.

In addition to the above hydrogen bonding within the major groove of the DNA, there are a number of apparent van der Waals contacts which are presumably also important in sequence-specific recognition. In particular, the C7 methyl groups of thymines +1, +3 and -5 appear to interact, respectively, with C^δ of Gln 27, with the carbonyl carbon of Asn 31 and the α -carbon of Lys 32 and with C^γ and C^ε of Lys 32 (Fig. 2). The importance of such hydrophobic interactions in the recognition of *lac* operator by lac repressor has been shown by Caruthers⁴¹.

Another feature of the model which may be of general relevance is the combination of sequence-specific interactions, mediated by the more rigid parts of the protein structure, together with non-specific interactions that are assumed to derive, in part, from flexibility in the cro structure. Two distinct types of flexible motion seem likely. First, the C-terminal residues (62–66) are not ordered in the crystals of the native protein, and are presumably free to move in solution. Because of this freedom of motion, and also because the C-terminal residues do not seem to be involved in sequence-specific interactions, we suggest that these residues are involved in the initial binding of the protein to the DNA. The second type of flexible

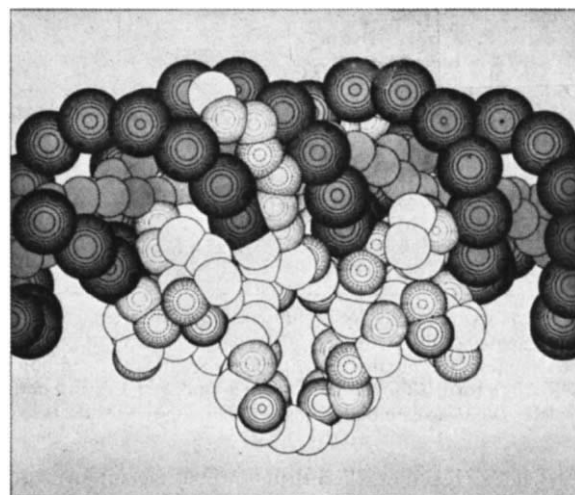


Fig. 5 Schematic drawing of the proposed sequence-specific complex of cro with DNA. For the stylized DNA (shaded) the large concentric circles indicate the positions of the phosphate groups and the smaller circles follow the bottom of the major and minor grooves. For the cro dimer, one circle is drawn for each amino acid; continuous concentric circles indicate acidic residues, broken circles indicate basic residues, dotted circles indicate non-charged hydrophilic residues and open circles show hydrophobic residues.

motion suggested by the cro structure is the 'hinge bending' motion described above. We suggest that such flexibility may be of general importance in facilitating 'sliding' along the DNA as has been proposed to occur for lac repressor⁴²⁻⁴⁷. Inspection of the structures of CAP² and cI repressor³ indicates that 'hinge' motion seems possible for these structures as well, suggesting that this type of flexibility may be of general significance. One could imagine that a hinge motion of the protein, as well as flexibility in the DNA, would give each of the cro monomers some freedom to 'sample' adjacent base pair atoms in the major groove of the DNA. Only at the correct twofold-symmetric binding site would the full DNA-binding specificity be satisfied for both cro subunits simultaneously. Because a dimer rather than a monomer of cro binds to DNA, the overall recognition site extends over 17 base pairs (~60 Å). However, the cro:DNA complex is presumably optimized for control rather than tightness of binding so that not all of the 17 base pairs need contact the protein and, in addition, departures from exact twofold symmetry of the base sequence could be used to

'fine tune' the strength of the cro:DNA complex.

Finally, we note that in the proposed cro:DNA complex, most of the sequence-specific interactions, as well as a number of the sequence-independent ones, come from the α_2 - α_3 helical unit. It has been striking to find, on the basis of recent sequence and structure comparisons, that a similar two-helical fold appears to occur in a number of other DNA-binding proteins⁴⁸⁻⁵⁴ indicating that the principles of specific protein-DNA recognition suggested by the proposed model may be of general significance.

We thank Alan Bender for help with model-building and Dr. P. H. von Hippel for discussions. This work was supported in part by an NIH postdoctoral fellowship to D.H.O. (GM08403), a Damon Runyon-Walter Winchell Cancer Fund postdoctoral fellowship to R.G.F. (DRG-388-FT), by the MRC of Canada through the Group on Protein Structure and Function (to W.F.A.), by NIH grants to B.W.M. (GM20066) and Y.T. (GM28138) and by grants to B.W.M. from the NSF (PCM-8014311) and the M. J. Murdock Charitable Trust.

Received 27 April; accepted 2 July 1982.

1. Anderson, W. F., Ohlendorf, D. H., Takeda, Y. & Matthews, B. W. *Nature* **290**, 754-758 (1981).
2. McKay, D. B. & Steitz, T. A. *Nature* **290**, 744-749 (1981).
3. Pabo, C. O. & Lewis, M. *Nature* **298**, 443-447 (1982).
4. Ptashne, M. *et al. Cell* **19**, 1-11 (1980).
5. Echols, H. in *The Molecular Genetics of Development* (eds Loomis, W. & Leighton, T.) 1-16 (Academic, New York, 1980).
6. Herskowitz, I. & Hagen, D. A. *Rev. Genet.* **14**, 399-445 (1980).
7. Maniatis, T. *et al. Cell* **5**, 109-113 (1975).
8. Watson, J. D. & Crick, F. H. C. *Nature* **171**, 737-738 (1953).
9. Arnott, S. & Hukins, D. W. L. *Biochem. biophys. Res. Commun.* **47**, 1504-1509 (1972).
10. Drew, H. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 2179-2183 (1981).
11. Peck, L. J. & Wang, J. C. *Nature* **292**, 375-378 (1981).
12. Rhodes, D. & Klug, A. *Nature* **292**, 378-390 (1981).
13. Bricogne, G. *Acta crystallogr.* **A30**, 395-405 (1974).
14. Molnar, C. E., Barry, C. D. & Rosenberger, F. U. *Tech. Memo. No. 229* (Computer Systems Laboratory, Washington University, St Louis, 1976).
15. Jack, A. & Levitt, M. *Acta crystallogr.* **A34**, 931-935 (1971).
16. von Hippel, P. H. in *Biological Regulation and Development* Vol. 1 (ed. Goldberger, R. F.) 279-347 (Plenum, New York, 1979).
17. Yarus, M. A. *Rev. Biochem.* **38**, 841-880 (1969).
18. Seeman, N. C., Rosenberg, J. M. & Rich, A. *Proc. natn. Acad. Sci. U.S.A.* **73**, 804-808 (1976).
19. Woodbury, C. P., Hagenbuchle & von Hippel, P. H. *J. biol. Chem.* **255**, 11534-11546 (1980).
20. Levitt, M. *J. molec. Biol.* **82**, 393-420 (1974).
21. Takeda, Y., Folkmanis, A. & Echols, H. *J. biol. Chem.* **252**, 6177-6183 (1977).
22. Johnson, A. D., Pabo, C. O. & Sauer, R. T. *Meth. Enzym.* **65**, 839-856 (1980).
23. Johnson, A. thesis, Harvard Univ. (1980).
24. Johnson, A., Meyer, B. J. & Ptashne, M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5061-5065 (1979).
25. Takeda, Y. *J. molec. Biol.* **127**, 177-191 (1979).
26. Wang, J. C. *Proc. natn. Acad. Sci. U.S.A.* **76**, 200-203 (1979).

27. Mandelkern, M., Elias, J. G., Eden, F. & Crothers, D. M. *J. molec. Biol.* **152**, 153-161 (1981).
28. Dickerson, R. E. & Drew, H. R. *J. molec. Biol.* **149**, 761-786 (1981).
29. Finch, J. T. *et al. Nature* **269**, 29-36 (1977).
30. Scheilman, J. A. *Biopolymers* **13**, 217-226 (1974).
31. Hatz, J. B., Magar, M. E. & Zimm, B. H. *Biopolymers*, **8**, 531-536 (1969).
32. Church, G. M., Sussman, J. L. & Kim, S.-H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1458-1462.
33. Zubay, G. & Doty, P. *J. molec. Biol.* **1**, 1-2 (1959).
34. Sung, M. T. & Dixon, G. H. *Proc. natn. Acad. Sci. U.S.A.* **67**, 1616-1623 (1970).
35. Adler, K. *et al. Nature* **237**, 322-327 (1972).
36. Warrant, R. W. & Kim, S. H. *Nature* **271**, 130-135 (1978).
37. von Hippel, P. H., Bear, D. G., Winter, R. B. & Berg, O. G. *Promoters: Structure and Function* (Eds Rodriguez, R. & Chamberlin, M.) (Praeger Scientific, New York, in the press).
38. Woodbury, C. P. & von Hippel, P. H. in *Gene Amplification and Analysis* Vol. 1 (ed. Chirikjian, J. C.) (Elsevier, New York, 1981).
39. Helene, C. *FEBS Lett.* **74**, 10-13 (1977).
40. Drew, H. R. & Dickerson, R. E. *J. molec. Biol.* **151**, 535-556 (1981).
41. Caruthers, M. H. *Accs Chem. Res.* **13**, 155-160 (1980).
42. Richter, R. H. & Eigen, M. *Biophys. Chem.* **2**, 255-263 (1974).
43. Berg, O. G. & Blomberg, C. *Biophys. Chem.* **4**, 367-381 (1976).
44. Berg, O. G., Winter, R. B. & von Hippel, P. H. *Biochemistry* **20**, 6929-6948 (1981).
45. Winter, R. B. & von Hippel, P. H. *Biochemistry* **20**, 6948-6960 (1981).
46. Winter, R. B., Berg, O. G. & von Hippel, P. H. *Biochemistry* **20**, 6961-6977 (1981).
47. Berg, O. G., Winter, R. B. & von Hippel, P. H. *Trends biochem. Sci.* **7**, 52-55 (1982).
48. Anderson, W. F., Takeda, Y., Ohlendorf, D. H. & Matthews, B. W. *J. molec. Biol.* (in the press).
49. Matthews, B. W., Ohlendorf, D. H., Anderson, W. F. & Takeda, Y. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1428-1432 (1982).
50. Steitz, T. A., Ohlendorf, D. H., McKay, D. B., Anderson, W. F. & Matthews, B. W. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3097-3100 (1982).
51. Ohlendorf, D. H., Anderson, W. F. & Matthews, B. W. *J. molec. Biol.* (submitted).
52. Sauer, R. T., Yocum, R. R., Doolittle, R. F., Lewis, M. & Pabo, C. O. *Nature* **298**, 447-451 (1982).
53. Weber, I. T., McKay, D. B. & Steitz, T. A. *Nucleic Acids Res.* (submitted).
54. Matthews, B. W., Ohlendorf, D. H., Anderson, W. F., Fisher, R. G. & Takeda, Y. *Cold Spring Harbor Symp. quant. Biol.* **47** (in the press).

Complete nucleotide sequence and identification of membrane components of the histidine transport operon of *S. typhimurium*

C. F. Higgins*, P. D. Haag, K. Nikaido, F. Ardeshir*, G. Garcia & G. Ferro-Luzzi Ames

Department of Biochemistry, University of California, Berkeley, California 94720, USA

The nucleotide sequence of the entire histidine transport operon from Salmonella typhimurium has been determined and is shown to consist of four genes, hisJ, hisQ, hisM and hisP. This operon provides the only example of a binding protein-dependent transport system for which the total number of protein components is known. Determination of the amino acid compositions and sequences of these four transport proteins, together with analysis of various transport mutants, allows us to propose a molecular model for binding protein-dependent transport.

At least five separate transport systems are involved in the uptake of histidine by *Salmonella typhimurium*¹. One of these, the high-affinity histidine permease, has been extensively

characterized and is among the best understood of all bacterial transport systems. This permease belongs to a broad group of 'shock-sensitive' transport systems that require for function specific substrate-binding proteins that are located in the periplasmic space and released by osmotic shock². Such periplasmic binding proteins are also involved in chemotaxis. Shock-sensi-

* Present addresses: Department of Biochemistry, University of Dundee, Dundee DD1 4HN, UK (C.F.H.); Department of Biochemistry, Stanford University, California 94305, USA (F.A.).

tive transport systems usually involve multiple components, in contrast to shock-resistant systems which require only a single, membrane-bound component (such as the lactose permease³). Previous studies of the high-affinity histidine permease have identified three essential genes, *hisJ*, *hisQ* and *hisP*, which together with a regulatory locus, *dhuA*, form a single operon located at 48.5 minutes on the *S. typhimurium* chromosomal map⁴ (Fig. 1). The *hisJ* gene encodes the periplasmic histidine-binding protein, J, which has been purified and characterized⁵. The *hisP* gene encodes a basic protein, P, which is located in the inner cytoplasmic membrane⁶. The previously defined *hisQ* gene is here shown to consist of two separate cistrons, *hisQ* and *hisM*. Thus, four protein components are required for high-affinity histidine transport.

The periplasmic protein J and the membrane-bound protein P interact with each other during histidine transport⁷. Complementation studies suggest that Q, M and P also interact with each other⁶, possibly existing as a multicomponent complex located in the cytoplasmic membrane. The Q, M and P proteins, but not the J protein, are also essential for an arginine transport system that requires an alternative periplasmic component, the lysine-arginine-ornithine-binding protein, LAO^{8,9}. The amino acid sequences of LAO and J are 70% homologous and the two proteins are believed to interact with the same site on the membrane-bound P protein during arginine or histidine transport, respectively¹⁰. Thus, the P protein provides a common membrane-bound receptor for two independent binding proteins. LAO is encoded by the *argT* gene which is located immediately upstream (Fig. 1) and is transcribed independently of the histidine transport operon¹¹.

Nucleotide sequence

The entire histidine transport operon has been cloned as a 12.4-kilobase (kb) *EcoRI* fragment in the phage vector λ gt4 (ref. 12). The precise locations of the transport genes on this fragment were determined by correlating accurately a restriction map with the genetic map of the histidine transport operon¹³. The various genes of the operon were sequenced by subcloning appropriate restriction fragments into the plasmid vectors pBR322 and pBR322-SVa¹³. Figure 1 shows schematically the histidine transport operon and the restriction fragments used in our study, while the complete nucleotide

sequence of the histidine transport operon, a total of 3,700 base pairs (bp), is presented in Fig. 2.

Identification of the coding sequences

***hisJ*.** The reading frame between bases 1 and 780 has previously been shown¹⁰ to be the coding sequence for *hisJ*. The amino acid sequence deduced from this nucleotide sequence corresponds with the known amino acid sequence of the J protein¹⁴. Salient features of the amino acid sequence of J and its relation to the LAO protein have been discussed elsewhere¹⁰.

***hisP*.** All *hisP* mutations are located within eight genetic regions (IX–XIIIb; see Fig. 1)^{4,6} (unpublished results). The structural gene encoding the P protein (as opposed to a possible regulatory region) must begin in or before region X because two-dimensional gel electrophoresis⁶ shows that one mutant mapping in this region (*hisP5566*, Fig. 1) results in a P protein with an altered mobility. As the *PstI* site is located in region XI¹³, the coding sequence for the P protein must include this site. The only open reading frame of any length spanning the *PstI* site is from position 2,367 to 3,140. This reading frame encodes a basic protein with a calculated molecular weight (M_r) of 28,738. The P protein has been confirmed experimentally to be basic and of ~24,000 M_r on SDS gels⁶; a discrepancy between the calculated molecular weight and the molecular weight determined on SDS gel is not unusual (see, for example, ref. 15). We confirmed that this reading frame encodes the P protein by purifying P protein, double-labelled with ³⁵S-methionine and ³H-leucine, from several two-dimensional gels and analysing the first 12 amino acids on a sequencer. ³H-leucine appeared in steps 5 and 10 (data not presented), and a very small amount of ³⁵S-methionine appeared in steps 1 and 2. These results are consistent with the predicted P sequence at the NH₂-terminus if we assume that most of the molecules have both N-terminal methionines removed post-translationally, but a few have only one or none of them removed.

***hisQ*.** The region between *hisJ* and *hisP* has previously been defined genetically as *hisQ* and divided into regions (I–VIII)⁶. The sequence (1,586 bp) between the termination codon of the J protein (position 781) and the initiation codon of the P protein (position 2,367) could potentially encode a single polypeptide of M_r 53,000. However, examination of the nucleotide sequence does not reveal a reading frame covering this entire region. We have identified a protein, which we call Q, by non-equilibrium two-dimensional gel electrophoresis¹⁶ either of whole cells or UV-irradiated cells¹⁷ following infection with λ phages carrying the cloned transport operon¹² (data not shown). It is a basic protein, of apparent M_r 21,000, located in the membrane, and present in extremely small amounts (10–100 times less than J; we have been unable to determine its NH₂-terminal amino acids). The Q protein is absent from mutants in regions I–IV, but present in mutants in regions V–VIII. Thus, the *hisQ* gene proper does not extend beyond region IV. Region I of *hisQ* contains the *HindIII* site b (Fig. 1)¹³; therefore, the *hisQ* coding sequence must start 5' to this restriction site. Of all the potential initiation codons (ATG, GTG and TTG¹⁸) between the end of *hisJ* and *HindIII* site b, only one, at position 966, could initiate a peptide of more than a few amino acids. This initiation codon is preceded by a strong ribosome-binding site and would initiate a basic protein of 24,573 M_r . Both the molecular weight and basicity of this protein are consistent with those of the Q protein. Therefore, we conclude that base pairs 966–1,649 are indeed the coding sequence for the Q protein and we re-define the *hisQ* gene as comprising this coding sequence only. Regions V–VIII (and part of IV) must code for one or more additional structural genes, or may comprise an extensive intergenic region.

***hisM*.** Between the termination codon for *hisQ* (position 1,650) and the beginning of *hisP* (position 2,367) is a sequence of position 717 which must correspond to regions V–VIII of the previously defined *hisQ* gene⁶. Mutants mapping in these regions are indistinguishable phenotypically from *hisP* and *hisQ*

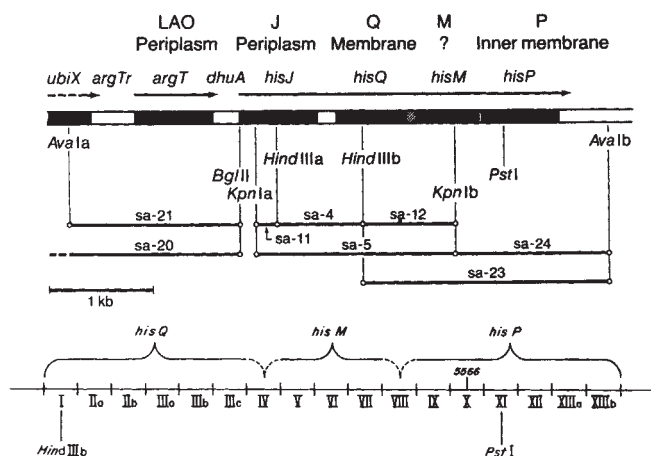


Fig. 1 Physical and genetic map of the histidine transport operon. Top: the structural genes (solid bars) and the intergenic and regulatory regions (open bars) are all drawn to scale and include the *argTr/argT* operon. The dashed bar indicates uncertainty about the initiation site for *hisM* (see text). Arrows above the genes indicate length and direction of transcripts, as deduced from the sequences. The transcription terminator sequence for the histidine transport operon is located 0.11 kb 3' to the end of *hisP*. The horizontal lines below the chromosome indicate specific restriction fragments subcloned for sequencing purposes. All are derived from the original clone of the entire region sa-1 (a 12.4-kb *EcoRI* fragment¹³). Bottom: genetic regions within the *hisQ*, *hisM* and *hisP* genes. These regions (in roman numerals) were defined by deletion mapping and biochemical analysis of mutants (refs 4, 6 and unpublished data).

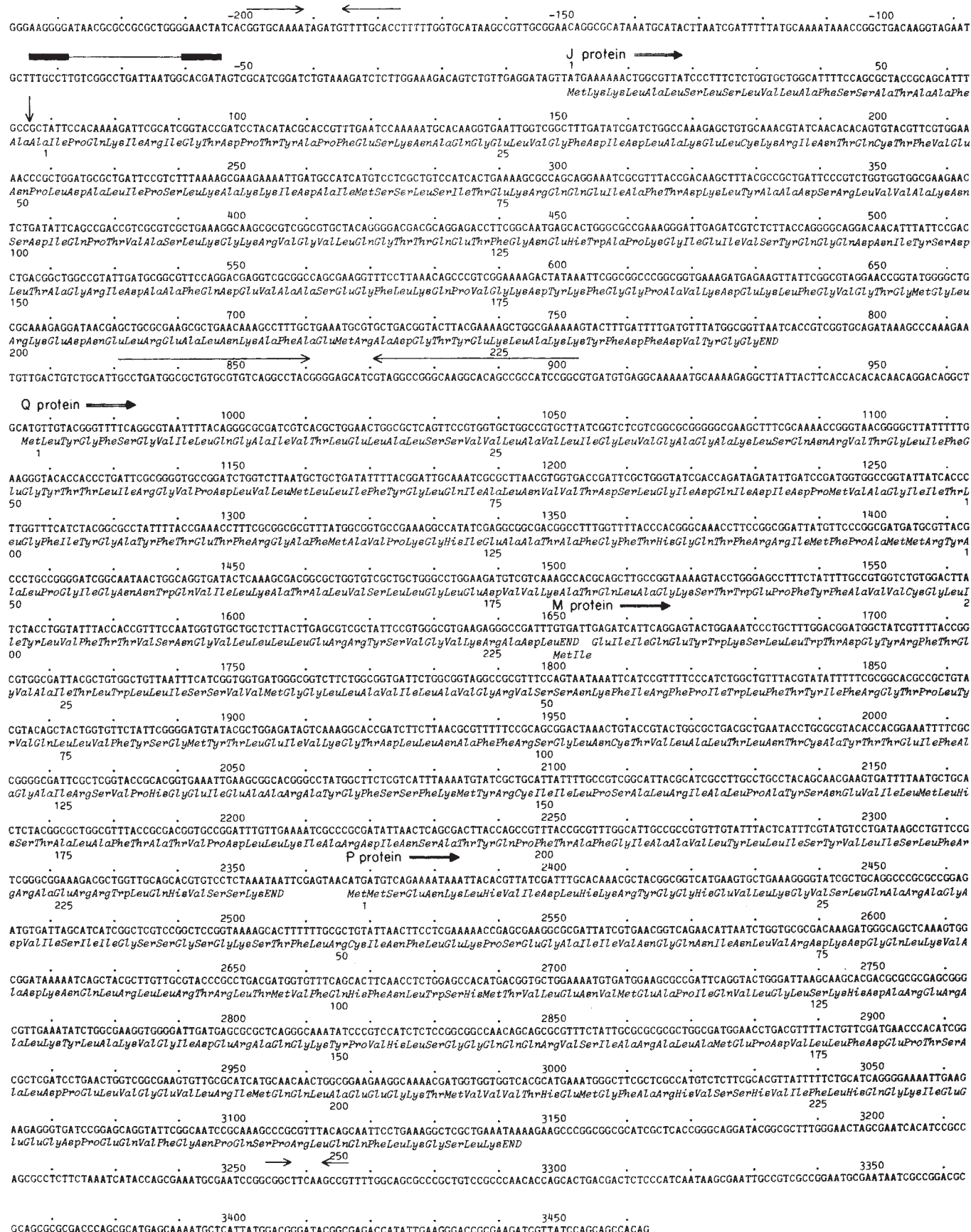


Fig. 2 Nucleotide sequence of the histidine transport operon and amino acid sequence of the gene products. The sequence presented (nonsense strand) begins immediately after the termination codon of the *argT* gene (see Fig. 1) and is numbered with base 1 being the first base of the *hisJ* gene; bases 5' to this initiation codon are numbered negatively. The amino acid sequences have been determined (J protein¹⁴) or predicted (all others); the P protein amino terminus has been confirmed by sequencing (see text). The promoter region in *dhuA* is marked by two solid bars above the -10 and -35 regions, joined by a line. The vertical arrow at position 66 indicates the site of processing of the J protein precursor. Horizontal arrows at bases 3,254-3,266 indicate the region of dyad symmetry which is the presumed transcription terminator signal for the transport operon. The horizontal arrows at -200 to -177 indicate a presumed transcription terminator signal for the *argT* gene which is immediately upstream from *dhuA*. Horizontal arrows at bases 831-902 indicate a region of dyad symmetry which is believed to serve a regulatory function²⁰. The *hisM* gene is assumed to start with the GTG at position 1,649, although initiation may be at the in-phase ATG at position 1,754 (see text). The entire nucleotide sequence was obtained by the method of Maxam and Gilbert⁴, both strands being sequenced at least once in their entirety. All restriction sites were sequenced over. The sequences of *dhuA* and *hisJ* have been published previously^{10,11}, but are presented here for completeness.

mutants and produce normal levels of unaltered P protein, indicating that this region is not regulatory and encodes at least one additional polypeptide. The nucleotide sequence of this region contains only one open reading frame of any length, beginning with either the GTG at position 1,649 or the ATG at 1,754. Each of these initiation codons is preceded by a ribosome-binding site. We call this sequence, between the GTG at 1,649 and the TAA translation termination codon at 2,353, *hisM*. This sequence encodes a basic protein of M_r 26,404, which we refer to as the M protein. This protein has not been identified, probably because of its hydrophobicity, strongly basic character and probable low level.

Noncoding regions

The region 240 bp 5' to the *hisJ* structural gene comprises the regulatory region for the histidine transport operon, *dhuA*. This region has previously been defined as the noncoding region between the termination codon of *argT* (position -240), the structural gene immediately upstream from the histidine transport operon, and the initiation codon (+1) of *hisJ*¹⁰. The promoter for the histidine transport operon occurs within the region -55 to -84 as indicated in Fig. 2. Features of the nucleotide sequence of *dhuA* important in regulating expression of the histidine transport operon have been discussed elsewhere¹¹.

Between *hisJ* and *hisQ* there is an intergenic region of 183 bp which contains an exceptionally long inverted repeat that, when transcribed, could form a very stable stem-loop structure ($\Delta G = -54 \text{ kcal mol}^{-1}$)¹⁹. As *hisJ* is expressed at a level at least 10 times that of *hisQ*, *hisM* and *hisP*, it seems probable that this intergenic element has a regulatory role in reducing the expression of the distal genes of the operon. The function of this element is discussed in the accompanying article²⁰.

There is no untranslated intergenic region between *hisQ* and *hisM* if the M protein starts with the GTG codon at position 1,649, as the first codon of *hisM* overlaps the termination codon of *hisQ*. However, if the initiation codon is the ATG at 1,754, the *hisQ*-*hisM* intergenic region consists of 101 bp: no inverted repeats or other features of potential regulatory interest have been found in this region. The third intergenic region, *hisM*-*hisP*, is only a few base pairs long. Such short intergenic regions, as seen for *hisQ*-*hisM* and *hisM*-*hisP*, are typical of multicistronic operons where the flanking genes are expressed to approximately the same extent^{21,22}.

Preceding each of the four genes of the histidine transport operon are presumed ribosome-binding sites, sequences homologous to the 3' end of 16S rRNA²³.

We have shown that the 5.4-kb *AvaI* fragment (Fig. 1) contains all the genes required for normal high-affinity histidine transport (data not presented). Every transport mutation we have isolated maps within this fragment¹³. Thus, if any histidine transport genes are located downstream from the *hisP* gene, they cannot be any larger than ~400 bp, this value being determined by *AvaI* site b (Fig. 1). However, no open reading frame of any length is found between the end of *hisP* and *AvaI* site b. A typical transcription termination signal (a G+C-rich dyad symmetry followed by several Ts; ref. 24) is found following *hisP* at positions 3,254-3,270. Thus, it is clear that we have identified all the genes required for high-affinity histidine transport.

Amino acid compositions of the histidine transport proteins

Figure 2 gives the predicted amino acid sequences of the four histidine transport proteins. The sequence of the J protein matches the determined amino acid sequence^{10,14}.

Of the four genes of the operon, only that for *hisJ*, the periplasmic histidine-binding protein, encodes a typical signal peptide (see examples in ref. 25), consisting of 22 mostly hydrophobic residues with two lysines located close to the NH₂-terminus. We have shown that the J protein is synthesized as a precursor, which is processed by removal of the signal peptide,

presumably during secretion into the periplasm (G.F.-L.A. and K.N., unpublished data). All other known periplasmic and outer-membrane proteins (including the *tra* proteins; E. G. Minkley, personal communication) are synthesized as precursors with a signal peptide.

The other three proteins of the operon do not have a typical NH₂-terminal signal peptide and indeed, for P we have shown that the NH₂-terminal sequence of the mature protein is the same as that predicted by the nucleotide sequence. We have previously shown P to be located in the inner (cytoplasmic) membrane; as it interacts with the J protein⁷, it is presumably exposed on the outside of the membrane. The Q protein is also located in either the inner or outer membrane (see above). It seems likely that the M protein also is located in the inner membrane because it is hydrophobic and the complementation pattern for mutants in these genes⁶ implies that all three proteins, P, Q and M, interact with each other. Thus, although an NH₂-terminal signal peptide seems to be essential for a protein to pass through the inner membrane, it is not required for insertion of proteins into the membrane. This is consistent with results obtained for several other inner membrane proteins whose genes have been sequenced and which also lack a typical NH₂-terminal signal peptide^{18,26}. However, there do seem to be certain exceptions to this general difference between inner membrane and outer membrane/periplasmic proteins. For example, the integral membrane protein bacteriorhodopsin from *Halobacterium halobium* is apparently synthesized as a precursor which is processed, although its structural gene does not code for a typical signal peptide^{27,28}; penicillin-binding proteins from *Escherichia coli* are also synthesized as precursors, though in this case it is not known whether their genes encode a typical signal peptide or whether processing occurs at the amino terminus²⁹. The lactose permease gene codes for what might possibly be a signal peptide, although no processing occurs^{3,30}. It is possible that in cases such as *hisM*, *hisQ* and *hisP* where no processing occurs, an internal hydrophobic region functions as a signal sequence, such as is found for the eukaryotic ovalbumin gene³¹.

Consideration of the amino acid composition of the four transport proteins shows that the Q, M and P proteins are basic, in contrast to the J protein and the bulk (~90%) of the bacterial cell proteins, which have *pI* values of about 7 or less (as estimated by two-dimensional gel electrophoresis of total cell protein³²). Two other inner membrane proteins, the lactose permease³ and NADH dehydrogenase¹⁸, are also basic, their basic character possibly being related to their need to interact with the negatively charged phospholipids of the membrane. Calculation of the hydrophobicity index³³ shows that proteins

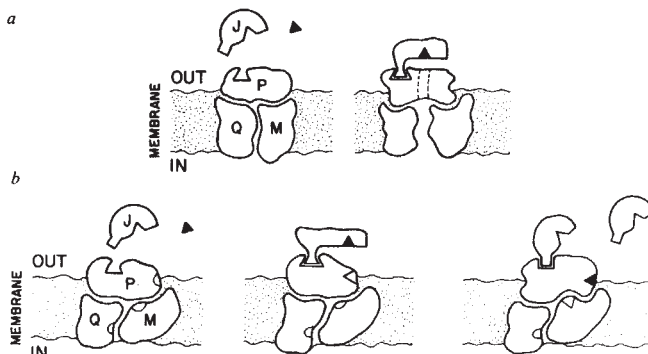


Fig. 3 Models for transport by a periplasmic system. *a*, 'Pore' model; *b*, 'substrate-binding' model. The solid triangle represents histidine. J is represented as interacting with P on having bound histidine. In model *a* a pore is formed through P (dashed lines) and between Q and M. However, it could be formed between P and/or Q and/or M molecules, and not through any single protein molecule. In model *b* potential histidine-binding sites on the membrane components are represented by hemispherical indentations which are sequentially 'activated' to functional sites (indicated by triangles). Only the initial transfer of histidine from J to P, with the concomitant conformational changes, is illustrated. The transfer of histidine from P to M and from M to Q can be imagined to occur by a similar series of events. See text for discussion of these models.

Q and M have several uncharged stretches of 10 or more amino acids (residues 62–80, 92–109 and 190–215 in Q; 20–47, 57–66, 68–85, 192–221 in M) which are more hydrophobic than water-soluble proteins³³. However, the P protein has no significantly hydrophobic stretches and in this respect is comparable with the water-soluble hydrophilic J protein. The percentage of charged amino acids in each of these proteins parallels the hydrophobicity calculations. The unusually low hydrophobicity of the membrane-bound P protein presumably reflects the fact that this protein interacts with the J protein⁷ and may therefore protrude extensively into the periplasm. Notably, histidine is found in all four transport proteins and is particularly abundant in the P protein (11 histidines). The predicted sequences of two histidine biosynthetic proteins, HisG and HisD, show that they also contain histidine (W. Barnes, personal communication). Thus, despite the fact that all these proteins serve to supply histidine to the cell, they also require histidine for their own synthesis. An analysis of codon usage in the four histidine transport genes shows clearly that codon usage is non-random, codons recognized by the major iso-accepting species of tRNA being most frequently used.

Mechanisms of transport

We have shown that the high-affinity histidine transport system requires the function of four separate proteins. One of these, the histidine-binding protein, J, is located in the periplasmic space. Of the other proteins, at least two (Q and P), and probably all three, are associated with the membrane. This provides the only example of a shock-sensitive, binding protein-dependent transport system for which the total number of essential components is known. However, all other binding protein-dependent systems which have been studied in detail also seem to require other components in addition to a binding protein^{34–41}. This multiplicity of components is in marked contrast to the lower-affinity, shock-resistant, binding protein-independent systems, such as the lactose permease (*lacY*), where only a single, membrane-bound component seems to be required.

The present determination of the number and location of the proteins required for transport, together with studies of transport mutants described previously and below, now allow us to begin to understand the molecular events involved in transport by a binding protein-dependent system (Fig. 3). The periplasmic protein J binds histidine⁵, inducing a conformational change (such a change has been shown to occur; ref. 42 and S. Zukin, personal communication), which then allows the J-histidine complex to interact with the membrane-bound protein P⁷. Proteins P, Q and M form a cytoplasmic membrane-bound complex⁶, with P exposed at the periplasmic face of the membrane. The membrane-bound complex of P, Q and M could promote the passage of histidine through the membrane either by forming a simple diffusion pore through the membrane, or they might possess specific binding sites for the transported substrate. These two possibilities are shown schematically in Fig. 3a and

b, respectively. As P, Q and M do not mediate transport to any significant extent in the absence of J, the interaction of the J-histidine complex with P must induce a conformational change in the membrane complex. In Fig. 3a the 'pore' model (modified from ref. 43) indicates that the J-histidine complex, on interaction with P, induces a conformational change allowing P, Q and M to form an open pore through which histidine can diffuse into the cell interior. In Fig. 3b, the 'binding site' model indicates that one or more of the three membrane-bound proteins has a histidine-binding site^{6,7}. These sites are 'activated' in a cascade-like manner, with histidine passing from one protein to the next, thus being carried across the membrane. It is also possible to envisage models which combine features from each of the two models described. In both models an energy-coupling step is presumably involved, although the mechanism by which this is achieved remains unknown.

The existence of certain classes of histidine transport mutant could help us distinguish between the two models. The wild-type histidine permease transports L-histidine efficiently; it also transports D-histidine and the histidine analogue HIPA with lower affinity. The isolation of mutants in *hisQ*, *hisM* or *hisP*, which differentially transport these substrates, suggests that the corresponding proteins have a substrate-binding site rather than being strictly part of a diffusion pore. We have found two types of such mutants: (1) *hisQ6699* transports L-histidine normally (K_m within 10% of normal) but does not transport D-histidine or HIPA at all within the limits of our measurements; (2) some *hisM* mutants (the 'odd group'⁴) result in the loss of transport of L- and D-histidine, and of HIPA, but also in an improved ability to transport L-histidinol. Both types of mutant can best be explained by an alteration in a substrate-binding site in a membrane-bound protein. This evidence supports the 'substrate-binding' model: isolation and characterization of additional mutants is required for confirmation. It has also been suggested that a membrane-bound component of the maltose permease may have a substrate-binding site (H. Shuman, personal communication).

The membrane-bound components of the histidine transport system are only present in very small amounts in the cell, and hence are extremely difficult to purify. The cloning and characterization of the genes encoding transport proteins thus provide the most powerful technique available for studying such systems. Given the many similarities between the histidine and other binding protein-dependent permeases, it seems likely that the model presented here can be generalized to most, if not all, such transport systems.

This work was supported by NIH grant AM12121 to G.F.-L.A., C.F.H. was supported by a NATO/SRC fellowship. P.D.H. was partially supported by a NSF predoctoral fellowship. We thank Dr R. Tjian and members of his laboratory for advice and for supplying us with ³²P-ATP, D. Cole and J. Sample for help with the sequenator experiment and W. Barnes for showing us DNA sequences of the histidine biosynthetic operon before their publication.

Received 9 February; accepted 16 June 1982.

- Ames, G. F.-L. in *1st ICN-UCLA Symp. molec. Biol.* (ed. Fox, C. F.) 409–426 (Academic, New York, 1972).
- Oxender, D. L. & Quay, S. C. in *Methods of Membrane Biology* Vol. 6 (ed. Korn, E. D.) 183–242 (Plenum, New York, 1976).
- Buchel, D. E., Gronenborn, B. & Muller-Hill, B. *Nature* **283**, 541–545 (1980).
- Ames, G. F.-L. *et al. J. Bact.* **129**, 1289–1297 (1977).
- Ames, G. F.-L. & Lever, J. E. *J. biol. Chem.* **247**, 4309–4316 (1972).
- Ames, G. F.-L. & Nikaido, K. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5447–5451 (1978).
- Ames, G. F.-L. & Spudich, E. N. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1887–1881 (1976).
- Kustu, S. G. & Ames, G. F.-L. *J. Bact.* **166**, 107–113 (1973).
- Kustu, S. G., McFarland, N. C., Hui, S. P., Esmond, B. & Ames, G. F.-L. *J. Bact.* **138**, 218–234 (1979).
- Higgins, C. F. & Ames, G. F.-L. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6038–6042 (1981).
- Higgins, C. F. & Ames, G. F.-L. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1083–1087 (1982).
- Ardeschir, F. & Ames, G. F.-L. *J. supramolec. Struct.* **13**, 117–130 (1980).
- Ardeschir, F., Higgins, C. F. & Ames, G. F.-L. *J. Bact.* **147**, 401–409 (1981).
- Hogg, R. W. *J. biol. Chem.* **256**, 1935–1939 (1981).
- Noel, D., Nikaido, K. & Ames, G. F.-L. *Biochemistry* **18**, 4159–4165 (1979).
- O'Farrell, P. Z., Goodman, H. M. & O'Farrell, P. H. *Cell* **12**, 1133–1142 (1977).
- Jaskunas, S. R., Lindahl, L. & Nomura, M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 6–10 (1975).
- Young, I. G., Rogers, B. L., Campbell, H. D., Jaworowski, A. & Shaw, D. C. *Eur. J. Biochem.* **116**, 165–170 (1981).
- Borer, P. N., Dengler, B. & Tinoco, I. *J. molec. Biol.* **86**, 845–853 (1974).
- Selker, E. & Yanofsky, C. *J. molec. Biol.* **130**, 135–143 (1979).
- Christie, G. E. & Platt, T. *J. molec. Biol.* **142**, 519–530 (1980).
- Shine, J. & Dalgarno, L. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1342–1346 (1974).
- Rosenberg, M. & Court, D. A. *Rev. Genet.* **13**, 319–353 (1979).
- Emr, S. D., Hall, M. N. & Silhavy, T. J. *J. Cell Biol.* **86**, 701–711 (1980).
- Gay, N. J. & Walker, J. E. *Nucleic Acids Res.* **9**, 2187–2194, 3919–3926 (1981).
- Chang, S. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 3398–3402 (1981).
- Dellwag, H. G. & Sumper, M. *FEBS Lett.* **116**, 303–309 (1980).
- Pratt, J. M., Holland, I. B. & Spratt, B. G. *Nature* **293**, 307–309 (1981).
- Ehring, R. D., Beyreuther, K., Wright, J. K. & Overath, P. *Nature* **283**, 537–540 (1980).
- Palmiter, R. D., Gagnon, J. & Walsh, K. A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 94–98 (1978).
- O'Farrell, P. H. *J. biol. Chem.* **250**, 4007–4021 (1975).
- Segrest, J. P. & Feldman, R. J. *J. molec. Biol.* **87**, 853–858 (1974).
- Randall, L. & Schwartz, M. *J. Bact.* **116**, 1436–1446 (1973).
- Shuman, H. A., Silhavy, T. J. & Beckwith, J. R. *J. biol. Chem.* **255**, 168–174 (1981).
- Bavofil, P., Hofnung, M. & Nikaido, H. *J. biol. Chem.* **255**, 8366–8369 (1980).
- Shuman, H. A. & Silhavy, T. J. *J. biol. Chem.* **256**, 560–562 (1981).
- Silhavy, T. J. *et al. Molec. gen. Genet.* **174**, 249–259 (1979).
- Anderson, J. J. & Oxender, D. L. *J. Bact.* **130**, 384–392 (1977).
- Lo, T. C. Y. *J. supramolec. Struct.* **7**, 463–480 (1977).
- Ordal, G. W. & Adler, J. *J. Bact.* **117**, 517–526 (1974).
- Robertson, D. E., Kroon, P. A. & Ho, C. *Biochemistry* **16**, 1443–1451 (1977).
- Singer, S. J. *et al. Rev. Biochem.* **43**, 805–833 (1974).
- Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).

LETTERS TO NATURE

Character of 'prolate' galaxies

Douglas O. Richstone & Michael D. Potter

Department of Astronomy, University of Michigan, Ann Arbor, Michigan 48109, USA

It is still not known whether elliptical galaxies are oblate, prolate or triaxial. The prototype 'prolate' galaxy NGC2685 has a rapid observed stellar rotation rate around the projected minor axis. We show here that a steady figure, exactly prolate model can have no such rotation, and that a nearly prolate model fails by a factor of 7. We suggest that the galaxy is oblate and that the gas disk is the wreck of a dwarf companion.

There has recently been much interest over the determination of the three-dimensional figure of elliptical galaxies. This interest can be classified into four areas—statistical studies of surface brightness¹⁻³, theoretical models^{4,5}, detailed surface photometry⁶⁻⁸ and the study of the distribution of embedded gas⁹. The last approach is distinguished by its potential ability to make statements about individual galaxies.

In the absence of figure rotation (for example, the tumbling of a bar) a gaseous disk is thought to settle into the fundamental plane of a biaxial galaxy in a time scale given by

$$\tau \approx T \left(6J_2 \cos \theta \left| \frac{d \ln T}{d \ln r} \right| \frac{\Delta r}{r} \right)^{-1} \quad (1)$$

where τ is the settling time, T is the orbital time, J_2 is the quadrupole moment of the gravitational potential, θ is the angle between the disk plane and the fundamental plane of the galaxy, r is the disk radius and Δr is the characteristic excursion in radius of gas clouds in the disk¹⁰. That excursion can be related to the peculiar velocity Δv of the gas clouds through $\Delta r = \Delta v / K$ where K is the epicycle frequency. Specializing now to a logarithmic potential (flat rotation curve), we can use $K = \sqrt{2}v_c/r$ (where v_c is the circular velocity) to write

$$\tau \approx T \left(3\sqrt{2}J_2 \cos \theta \frac{\Delta v}{v_c} \right)^{-1} \quad (2)$$

Noting that J_2 is <0.5 even for very flat galaxies and that $\Delta v/v_c \approx 0.1$ (a reasonable assumption), we see that the settling time from an inclination of 60° or more is always $>\tau_0 \sim 9T$.

The fact that the characteristic settling time is less than the age of the Universe has been used to argue that NGC2685¹¹ is, in fact, a prolate object with a small pattern speed¹².

The next step is to construct a prolate stellar dynamical model for NGC2685 which is consistent with the observed stellar rotation. We attempted such a construction as part of a systematic study of prolate stellar systems using Schwarzschild's⁴ self-consistent field method. That method chooses a specific potential, follows stellar orbits in that potential and finally adds up the time averaged stellar orbits to produce a density distribution corresponding to the original potential. We used a logarithmic potential with a flattening corresponding to an E5 galaxy⁵.

Such a model must reproduce the large observed rotation rate about the projected minor axis. The observed rotation parameter v/σ is ~ 1 (ref. 13), while σ is $\sim 100 \text{ km s}^{-1}$. (This dimensionless rotation rate is 3–10 times the normal observed value for an elliptical, and is much more consistent with an S0 disk). There is some evidence also for minor axis notation. If this is true, it would suggest that the galaxy is triaxial.

Because the figure is fixed, the observed rotation must be derived from orbits which stream through the figure. Such an orbit is shown in Fig. 1a. Maximizing the polar streaming velocity (using linear programming techniques) permits a maximum streaming velocity—corresponding to an observed v/σ of only 0.075 on the projected major axis. Moreover, this

large streaming velocity can only be obtained if the y component l_y of the total angular momentum of each orbit is >0 for all $l_z = 0$ orbits (no orbit circulates in the opposite sense). Because these orbits are only neutrally stable against perturbations which cause rapid precession about the z axis, this configuration is implausible. Even if the orbits were stable, the model's v/σ completely fails to reproduce the observed one of unity.

A model more consistent with the observations can be constructed with a slightly triaxial figure. The orbit of Fig. 1a can easily be followed in a gravitational potential of the form

$$\Phi = \frac{1}{2} \ln \left(x^2 + \frac{y^2}{p^2} + \frac{z^2}{q^2} \right) \quad (3)$$

If $p < 1$ the orbit is stable but l_z is no longer conserved and the orbit is no longer required to reach $z = 0$ frequently (see Fig. 1b). Therefore a larger number of these orbits can be incorporated into the model. A nearly prolate model was constructed maximizing the influence of these streaming orbits with $p = 0.9$. Viewed from the x axis (edge on) its $v/\sigma = 0.1$. The maximum v/σ was achieved from $\sim 30^\circ$ from the x axis in the x - z plane. From that vantage point $v/\sigma = 0.14$. Thus we cannot construct a nearly prolate model with no figure motion and large v/σ . We emphasize that this method maximizes v/σ within the set of all self-similar solutions of this general form.

The failure to construct a figure fixed prolate model with large v/σ indicates that the galaxy is not prolate. An alternative explanation for the gas ring is that it is the wreck of a dwarf

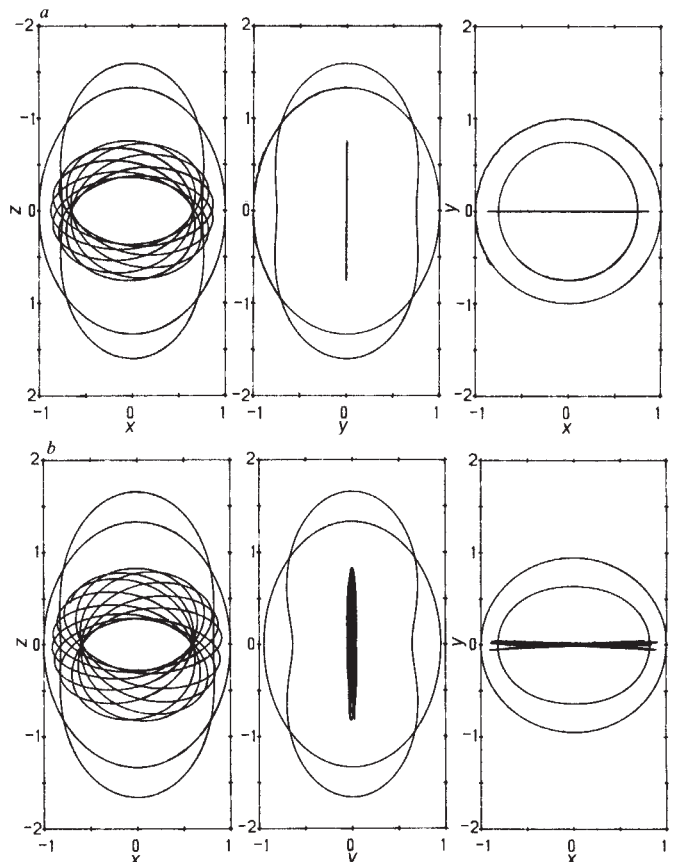


Fig. 1 a, Three orthogonal views of an orbit in a fixed biaxial prolate potential (with $p = 1$, $q = 1.333$, see equation (3)). The envelopes are cuts through the equipotential and equal-density surfaces. Note that the orbit streams through the galaxy maintaining a well defined circulation in the x - z plane. This orbit has no z angular momentum and is only neutrally stable. b, As a but with a triaxial potential with $p = 0.95$ and $q = 1.333$. This orbit also streams. It is a stable orbit with fluctuating z angular momentum.

companion, which has not yet decayed to the fundamental plane of the galaxy.

Such a scenario does not, by itself, constrain the age of this gas disk. The disk is very inclined, and equation (2) suggests that for exact polar alignment ($\theta = 90^\circ$), the disk lifetime is infinite.

In fact, equation (2) for the settling time also provides the means to calculate the time independent (or ensemble average) distribution of ring angles. From equation (2), $d\mu/dt \propto \mu^2$ where $\mu = \cos \theta$ is the cosine of the tilt angle ($\mu = 0$ for a polar ring). Because rings are inserted at some angle and then decay to the plane, they obey a continuity equation of the form

$$\frac{\partial n}{\partial t} + \frac{\partial}{\partial \mu} \left(\frac{d\mu}{dt} n \right) = \left(\frac{\partial n}{\partial t} \right)_i \quad (4)$$

where $n = n(\mu, t)$ is the chance of finding a disk at inclination μ at time t , and $(\partial n / \partial t)_i$ is the insertion rate. In steady state, $\partial n / \partial t = 0$. For insertion in strictly polar orbits it is easy to show that $n \propto \mu^{-2}$, while for an insertion rate independent of μ , $n \propto \mu^{-1}$ (to verify these results it is necessary to apply the appropriate boundary conditions at $\mu = 0$).

Note that the above results indicate a preference for disks in non-steady state but relatively long-lived nearly polar configurations regardless of the injection rate. The number of inclined disks will, however, be set by the ratio of the injection rate to the decay rate. Since the latter is $\approx \tau_0 \approx 10T$ ($\approx 10^9$ yr), a frequent incidence of tilted disks is required to demand an insertion time scale much shorter than the age of galaxies. However, these objects may be more common than we realize judging from a flurry of recent discoveries (ref. 14 and P. Schechter, unpublished data). Note that our Galaxy may be about to acquire such rings as the Magellanic clouds come apart¹⁵.

Although we have presented a strong case for the view that this particular galaxy is axisymmetric with two long axes, we have not explicitly ruled out a slowly tumbling bar.

There is, however, an argument which suggests that slow figure motion does not alter the small values of v/σ derived above. Schwarzschild¹⁶ and Miller and Smith¹⁷ have constructed models of tumbling triaxial and biaxial prolate galaxies of very different rotation rates. In both studies the observable rotational velocity is never more than 1.5 times the pattern speed at any radius. Because the pattern speed must be small to avoid disrupting the polar rings in the galaxy under consideration if it is a tumbling bar¹², these models also indicate a small value of v/σ .

We conclude that it is easy to model NGC2685 as an oblate galaxy with polar rings of gas which are the wreck of a dwarf companion. It is impossible to construct a fixed figure prolate model which reproduces the observed stellar rotation rate. It is doubtful that a tumbling prolate model would resolve the discrepancy. It may be that a number of galaxies which have been identified as prolate objects because of embedded disks are in fact oblate objects.

We thank G. Lake, S. Tremaine, M. Schwarzschild and P. Schechter for advice and encouragement. This research was partially supported by NSF grant AST 82-02930 and by NASA grant NGR 23-005464.

Received 24 May; accepted 24 June 1982.

1. Marchant, A. B. & Olson, D. W. *Astrophys. J. Lett.* **230**, L157-L159 (1979).
2. Richstone, D. O. *Astrophys. J.* **234**, 825-828 (1979).
3. Olson, D. W. & de Vaucouleurs, G. *Astrophys. J.* **249**, 68-75 (1981).
4. Schwarzschild, M. *Astrophys. J.* **232**, 236-247 (1979).
5. Richstone, D. O. *Astrophys. J.* **238**, 103-110 (1980).
6. Williams, T. B. & Schwarzschild, M. *Astrophys. J.* **227**, 56-63 (1979).
7. Williams, T. B. & Schwarzschild, M. *Astrophys. J. Suppl.* **41**, 209-213 (1979).
8. Bertola, F. & Galletta, G. *Astr. Astrophys.* **77**, 363-365 (1979).
9. Bertola, F. & Galletta, G. *Astrophys. J. Lett.* **226**, L115-L118 (1978).
10. Tohline, J. E., Simonson, G. F. & Caldwell, N. *Astrophys. J.* **252**, 92-101 (1981).
11. Sandage, A. *The Hubble Atlas of Galaxies*, 7 (Carnegie Institution, Washington DC, 1961).
12. Tohline, J. E. & Durisen, R. H. *Astrophys. J.* **257**, 94-102 (1982).
13. Schechter, P. L. & Gunn, J. E. *Astr. J.* **83**, 1360-1362 (1978).
14. Hawarden, T. G., Elson, R. A. W., Longmore, A. J., Tritton, S. B. & Corwin, H. G. *Mon. Not. R. astr. Soc.* **196**, 747-750 (1981).
15. Tremaine, S. D. *Astrophys. J.* **203**, 72-74 (1976).
16. Schwarzschild, M. Preprint (Princeton University, 1982).
17. Miller, R. H. & Smith, B. F. *Astrophys. J.* **227**, 785-797 (1979).

Composition-dependent glass-transition temperatures and copolymers

P. R. Couchman

Department of Mechanics and Materials Science, Rutgers University, PO Box 909, Piscataway, New Jersey 08854, USA

The glass transition is of primary relevance to the forming conditions and end-use properties of polymeric materials. Several polymers of industrial importance, most notably polyvinyl chloride (PVC), contain small molecules as processing lubricants (plasticizers) and are, therefore, solutions. Separately, economic and materials advantages have spurred the use of one-phase mixtures (miscible blends) of high polymers. The general theme of the compositional variation of glass-transition temperatures (T_g) includes, in addition, the variation of T_g with degree of polymerization, with cross-link density (network formation) and with copolymer composition. An attempt is made here to find a unified basis for the prediction of these several aspects of an identifiably single topic.

For solutions which are largely random and for phenomena which can be modelled in terms of such solutions, a recent thermodynamic theory for the compositional variation of glass-transition temperatures¹ has provided a method for calculating solution glass-transition temperatures, T_g , from pure-component glass-transition temperatures, T_{gi} , their corresponding glass-transition increments of heat capacity, ΔC_{pi} , and the mole or mass fractions, X_i , of these components. The equation for the glass-transition temperature of a c -component blend is, in the temperature-independent approximation for the ΔC_{pi} ,

$$\ln T_g = \frac{\sum_{i=1}^c X_i \Delta C_{pi} \ln T_{gi}}{\sum_{i=1}^c X_i \Delta C_{pi}} \quad (1)$$

Systems to which this entropic expression has been applied successfully include miscible blends of polymers¹⁻³ and plasticized straight-chain⁴ and network⁵ polymers. Additionally, the effect of degree of polymerization⁶ and of cross-link density⁷ on T_g have been accounted for by the theory.

For binary solutions, the derivative

$$\frac{d \ln T_g}{dX_2} = \frac{\Delta C_{p1} \Delta C_{p2} \ln T_{g2}/T_{g1}}{(X_1 \Delta C_{p1} + X_2 \Delta C_{p2})^2} \quad (2)$$

does not vanish (except in the trivial case) and, therefore, equation (1) predicts a monotonic dependence of T_g on overall composition. For near-random (and other) copolymers T_g can, however, manifest a compositional absolute extreme value. Thus, any model of copolymers as binary solutions cannot be generally acceptable. Nevertheless, as outlined below, it is possible to include copolymers of arbitrary ordering within the theory by recognizing formally that when differences between primary bonds must be considered in connection with the glass-transition properties of entities comprising a solution, the microscopic rather than overall composition determines the solution behaviour.

Sequence-distribution effects on T_g can be accounted for in the nearest-neighbour approximation by formal treatment of the general copolymer as a random solution of mer pairs, fractions f_{ii} , f_{ij} , f_{ji} . The particular version of equation (1) which then arises is

$$\ln T_g = \frac{f_{ii} \Delta C_{p11} \ln T_{g11} + f_{ij} \Delta C_{p12} \ln T_{g12} + f_{ji} \Delta C_{p21} \ln T_{g21}}{f_{ii} \Delta C_{p11} + f_{ij} \Delta C_{p12} + f_{ji} \Delta C_{p21}} \quad (3)$$

The denominator of this relation is, as for equation (1), the solution transition increment of heat capacity, ΔC_p (ref. 1). For copolymers formed by continuous reaction of i and j monomers

at constant feed composition, the dyad fractions of equation (3) can be rewritten as a function of the fraction, f_i , of i monomers in the feed and of the monomer reactivity ratios⁸, r_i , r_j (these are a measure of the relative rates at which a chain end grows by the addition of like and unlike mers). The consequent relation,

$$\ln T_g = \frac{r_i f_i^2 \Delta C_{pii} \ln T_{gii} + 2f_i(1-f_i) \Delta C_{pij} \ln T_{gij} + r_j(1-f_i)^2 \Delta C_{pjj} \ln T_{gjj}}{r_i f_i^2 \Delta C_{pii} + 2f_i(1-f_i) \Delta C_{pij} + r_j(1-f_i)^2 \Delta C_{pjj}} \quad (4)$$

provides a formal method for the calculation of glass-transition temperatures of the quasi-binary solution from glass-transition properties of the corresponding homopolymers, the fully alternating copolymer, and the monomer reactivity ratios. For values of these properties which admit a physically acceptable value of f_i as a solution of the quadratic equation obtained from the condition $d \ln T_g / df_i = 0$, the copolymer glass-transition temperature will pass through an absolute extremum. The sign of the second derivative $d^2 \ln T_g / df_i^2$ at this value of f_i determines the nature of the extremum. If the quadratic equation gives no admissible value of f_i , the copolymer glass-transition temperature is a monotonic function of feed composition. The compositional variation of the copolymer glass-transition increment of heat capacity can be analysed in a similar manner. Generally, extreme values of T_g and ΔC_p occur at different compositions.

A considerable simplification of the theory occurs for an ideal azeotropic copolymerization ($r_i = r_j = 1$) of monomers for which $T_{gii} = T_{gjj}$ and $\Delta C_{pii} = \Delta C_{pjj}$. In these circumstances, the extreme value of the copolymer glass-transition temperature, T_g^* , occurs at $f_i = \frac{1}{2}$ and is given by

$$\ln T_g^* = \frac{\Delta C_{pii} \ln T_{gii} + \Delta C_{pji} \ln T_{gji}}{\Delta C_{pii} + \Delta C_{pji}} \quad (5)$$

If, further, the difference between ΔC_{pii} and ΔC_{pji} can be neglected,

$$T_g^* = (T_{gii} T_{gji})^{1/2} \quad (6)$$

Comparison of theory and experiment for ethylacrylate-vinylidene chloride copolymers affords a critical test of the theory, as a well-defined extreme value of T_g occurs. For this system⁹, $r_i \approx r_j \approx 1$, $T_{gii} \approx T_{gjj} \approx 250$ K and $T_{gij} = 408$ K. In the absence of information on the dyad transition increments of heat capacity, the approximation $\Delta C_{pii} = \Delta C_{pji} = \Delta C_{pji}$ allows equation (6) to be used, predicting a maximum value of the copolymer glass-transition temperature of 319 K, at $f_i = \frac{1}{2}$. The observed maximum value of T_g is 308 K and occurs at $f_i \approx \frac{1}{2}$ (ref. 9).

Previous applications of the theory¹⁶ have placed in a more general context a variety of empirical and semi-empirical expressions for specific aspects of the compositional variation of T_g . Similarly, the formal treatment outlined here gives as first or second-order approximations current relations for the variation of copolymer glass-transition temperatures with composition. Finally, note that equations (1) and (3) are in effect integrated, predictive forms of an entropic formal tautology for composition-dependent transitions of the second-order kind^{10,11}.

This work was supported by the NSF, Polymers Program.

Received 15 April; accepted 24 June 1982.

Gold deposition at low temperature on sulphide minerals

G. Michael Bancroft & Gilles Jean

Department of Chemistry, University of Western Ontario, London, Ontario, Canada N6A 5B7

Gold deposits are commonly associated with quartz veins containing variable amounts of pyrite and other sulphides. Gold is thought to be transported as AuCl_2^- or AuCl_4^- in strongly acidic and saline solutions and as AuS^- or $\text{Au}(\text{S}_2\text{O}_3)_2^{3-}$ in weakly acidic and alkaline solutions by hydrothermal solutions coming from great depth and rising in rock formations through faults^{1,2}. In some cases gold occurs as finely divided particles, intimately intergrown with pyrite, suggesting that gold, quartz and pyrite were probably precipitated in response to a change in temperature. Fyfe and Henley³ showed that the solubility of gold increases from 10 p.p.m. at 300 °C to 1,000 p.p.m. at 510 °C. In other cases gold may be aggregated in one spot in abundance while the remainder of the ore is barren, indicating preferential precipitation in one site. Using ESCA (electron spectroscopy for chemical analysis), we have studied the reaction between KAuCl_4 in solution and various sulphide minerals. Gold(III) was rapidly reduced to metallic gold on the surface of these minerals, suggesting a new mechanism for the deposition of gold at low solution concentration and temperature.

Various sulphide minerals [pyrite (FeS_2), pyrrhotite (FeS), sphalerite (ZnS), galena (PbS)] were cut into plates of $1 \times 1 \times 1$ cm and were reacted with KAuCl_4 solutions of 1, 10 and 100 p.p.m. at different pH for various periods of time. Figure 1a shows the ESCA spectrum of the gold 4f lines of the original KAuCl_4 solution evaporated on Teflon. The binding energy of the gold 4f_{7/2} line at ≈ 87.4 eV (relative to carbon 1s line at 284.5 eV) is representative of Au(III) compounds⁴. Figure 1d is the spectrum of pure metallic gold and shows a binding energy for the gold 4f_{7/2} line of ≈ 83.5 eV which is 3.9 eV lower than Au(III). Figure 1b shows the ESCA spectrum of sphalerite (ZnS) reacted with a solution of KAuCl_4 (10 p.p.m.) at room temperature. This spectrum clearly shows that both Au(III) and Au(0) are present on the surface. Figure 1c shows the ESCA spectrum of pyrite (FeS_2) reacted with

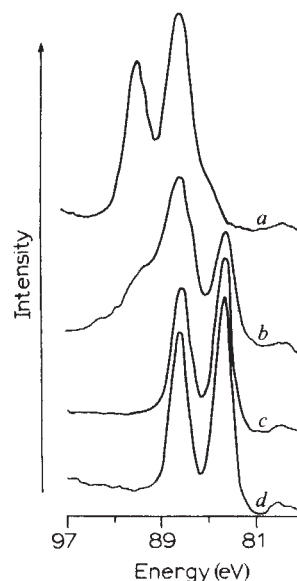


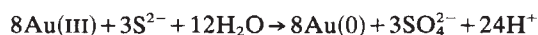
Fig. 1 Gold 4f ESCA spectra for: a, KAuCl_4 solution (10 p.p.m.) evaporated onto Teflon; b, sphalerite (ZnS) plate reacted with this KAuCl_4 solution; c, pyrite (FeS_2) plate reacted with this KAuCl_4 solution; d, metallic Au.

1. Couchman, P. R. *Macromolecules* **13**, 1272-1276 (1980).
2. Leisz, D. M., Kleiner, L. W. & Gertenbach, P. G. *Thermochim. Acta* **35**, 51-58 (1980).
3. Ryan, C. L. thesis, Univ. Massachusetts (1981).
4. Bair, H. E. & Warren, P. C. *J. macromolec. Sci.-Phys.* **B20**, 381-402 (1981).
5. Ellis, T., Karasz, F. E. & Moy, P. *Polym. Preprints* **22**, 121-122 (1981).
6. Couchman, P. R. *J. mater. Sci.* **15**, 1680-1683 (1980).
7. Martin, C. G., Mehta, R. K. & Lott, S. E. *Polym. Preprints* **22**, 319-320 (1981).
8. Wall, F. T. *J. Am. chem. Soc.* **66**, 2050-2057 (1944).
9. Comyn, J. & Fernandez, R. A. *Eur. Polym. J.* **11**, 149-151 (1975).
10. Landau, L. & Lifschitz, E. M. *Statistical Physics* 2nd edn (Pergamon, Oxford, 1969).
11. Angell, C. A., Sare, J. M. & Sare, E. J. *J. phys. Chem.* **82**, 2622-2629 (1978).

KAuCl₄ in solution. The binding energy of ≈ 83.5 eV indicates that Au(0) only is present on the surface.

The reduction of Au(III) was found to be independent of either the pH or the initial concentration of gold and was found to occur even in saturated KCl solution. Gold was reduced from Au(III) to Au(0) within a minute by every mineral studied except sphalerite. In this latter case, the reduction was slower since both species, Au(III) and Au(0), were observed. This suggests that the nature of metal ion in the sulphide may have an important role in the deposition of gold. Further work is being done to clarify this point.

The mechanism that is thought to occur is: adsorption of Au(III) as the hydrated or the hydrolysed species (chloride is not seen on the surface), followed by reduction of the gold by the sulphide.



The deposition of gold goes on even after the surface is entirely covered, and can actually be seen with the naked eye. After the first layer of Au(0) is deposited, electrons for the reduction are probably coming from the interior of the crystal. This phenomenon is similar to the observed reduction of gold on activated carbon⁵. These results suggest that adsorption of Au(III) by sulphide minerals followed by reduction could have an important role in the deposition of gold in natural systems, especially at low temperature and low solution concentrations of gold.

We thank the NSERC (Canada) for research funds, Professor R. J. Puddephatt for the KAuCl₄, and Professor W. S. Fyfe for helpful discussions.

Received 29 March; accepted 28 June 1982.

1. Wedepohl, K. H. in *Handbook of Geochemistry* Vol. II-S, 79.F.1 (Springer, Berlin, 1978).
2. Goleva, G. A., Krivenkov, V. A. & Gutz, Z. G. *Geochim. Int.* **7**, 518–529 (1970).
3. Fyfe, W. S. & Henley, R. W. *Miner. Sci. Engng* **5**, 295–303 (1973).
4. Carlson T. A. *Photoelectron and Auger Spectroscopy* (Plenum, New York, 1975).
5. McDougall, G. J. & Hancock, D. *Gold Bull.* **14**, 138–153 (1981).

Layered convection and crystal layers in multicomponent systems

R. C. Kerr* & J. S. Turner

Research School of Earth Sciences, The Australian National University, Box 4, Canberra 2600, Australia

Laboratory experiments^{1–5} in aqueous solutions have been used to model some of the physical and dynamical processes occurring during crystallization in magma chambers. In particular, it has been demonstrated^{2,3,5} that crystallization at the side walls can lead to density stratification and layering in an initially homogeneous container, and also to the formation of vertical compositional gradients (that is, to 'differentiation') when more than one solute is present. Density gradients and compositional layers can alternatively be set up by replenishing a container from below with fluid of different temperature and composition^{4,6,7}. Underlying all this work has been the concept that some aspects of the regularity and large horizontal extent of layering in well-documented igneous intrusions^{8,9} might be explained in terms of 'double-diffusive' effects in the fluid state, with the growth of crystals subsequently being affected by these pre-existing layers. In the latest experiments reported here, we have explicitly addressed the problems of horizontal layering produced in the crystals themselves, when a compositionally-stratified tank of fluid containing several solutes has been cooled from either below or above. The relationship between the scales of the layers in the fluid and in the crystals has also been examined.

* Present address: Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW, UK.

Earlier experiments^{1,4} in which crystals of a constant composition were grown at the top or bottom of an experimental tank were of two kinds, in both of which layered structures were produced. When an aqueous solution of a single solute (usually Na₂CO₃) was cooled from the top¹, crystals of Na₂CO₃ · 10H₂O formed in horizontal layers, whether the initial fluid was homogeneous, was smoothly stratified, or was set up with several distinct layers of different densities. The dominant effect in all cases was the formation of a cold but less dense fluid layer against the cooled upper boundary as the denser crystals grew (because the density of the saturated solution decreased as the temperature fell). When the top boundary was held below the eutectic temperature, bands were observed in the eutectic layer at levels corresponding to pre-existing fluid interfaces; the advancing front of the crystals growing below this also remained nearly horizontal, and the growth rate was clearly influenced by the positions of the fluid layers. Cooling a compositional gradient from below produced a compact layer of dense crystals of nearly uniform thickness, above which was a fluid layer which increased in depth with time. In this layer convection was driven by the compositional change in the residual fluid generated by the removal of the solid phase at the lower boundary, and the changing depth was a measure of the total volume of less dense fluid produced.

The second type of experiment involving bottom crystallization was conducted⁴ to model the evolution of a magma chamber following the sudden influx of a new layer of hot but heavy magma at the base. This lower layer remained chemically isolated from the rest of the chamber, while it cooled rapidly through the interface and crystallization proceeded⁶. In the laboratory model of this process, the lower layer was hot KNO₃ solution and the upper layer was a colder, less dense NaNO₃ or K₂CO₃ solution. Crystallization of KNO₃ took place on the bottom, often extending upwards to the interface, which produced a sharp horizontal top to the crystal layer. At the same time the density of the residual liquid in the lower layer decreased due to the compositional change. When this density became equal to that of the upper layer, the interface broke down, so that the two layers mixed thoroughly together leaving a layer of KNO₃ crystals at the base. Related experiments⁷ have been conducted with a compositional gradient above the lower layer, and these show that the overturned lower-layer fluid is then confined by the gradient to the lower part of the upper region.

The present experiments represent extensions of each of these cases, with the additional feature that more than one component crystallized out during the course of an experiment, in a sequence and manner which were determined by the evolution of the layered fluid system. Examples have been chosen to illustrate different ways in which crystallization can be controlled by the compositional stratification.

The experimental tank used was a Perspex cube of 20 cm internal dimensions, fitted with a metal base through which coolant could be circulated at a controlled temperature. The procedure was to fill the tank with one or more layers of solution of known composition and temperature (and hence density), to insulate all the side boundaries, and to monitor at intervals the development of the fluid and crystal layers, and in some cases the properties of the residual fluid. Usually the bottom was cooled slowly, by lowering the temperature of the coolant in small steps, but in several experiments the bottom crystallization was produced by cooling from the top, followed by internal heat transfer (through the interfaces between the layers).

Consider first an experiment in which the whole tank was filled with a homogeneously mixed aqueous solution of CuSO₄ and Na₂SO₄. The initial compositions were ~ 19 wt% CuSO₄ and 9 wt% Na₂SO₄, well in the range where the copper salt will crystallize first. (The conditions used in this and later experiments are summarized in Fig. 1, where the four cases a–d correspond to the photographs in Fig. 3.) When this solution was cooled slowly through the bottom of the tank, a layer of CuSO₄ · 5H₂O crystals did indeed grow, to a thickness of ~ 1 cm.

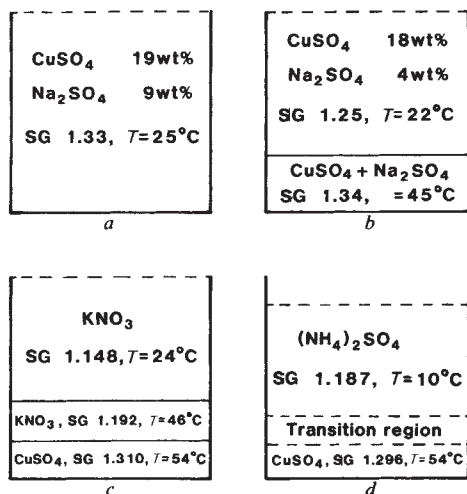


Fig. 1 Summary of the initial conditions in the four experiments for which the final crystal products are shown in Fig. 3a–d. The tank was cooled through the base in each case, and insulated at the sides and the top.

At the same time the composition of the residual fluid, well-mixed by 'compositional convection', moved towards the cotectic line as shown on Fig. 2, which is part of a phase diagram prepared for a more detailed study⁵ of crystallization of these two components in a different geometry. Before the cotectic line was reached, however, Na₂SO₄·10H₂O crystals started growing on top of the previous layer (with just a few crystals of CuSO₄·5H₂O visible between them), and built up to a nearly uniform depth of 2 cm, as shown in Fig. 3a, while the composition evolved further back into the copper field. This behaviour has been attributed⁵ to disequilibrium effects, and the different kinetics of crystallization of the two components, which made it possible for the growth of the copper salt to be suppressed when it was covered by the faster-growing Na₂SO₄·10H₂O. Similar kinetic effects are also likely in crystallizing magmas, but even if these remained much closer to the equilibrium conditions, the evolution towards, and then along, a cotectic line would nevertheless produce an abrupt change in composition of the crystallizing phases.

The next experiment to be discussed (Fig. 1b) started with a 6-cm deep layer of similar composition to that used in the preceding experiment, with specific gravity 1.34 and temperature 45°C, underlying a layer of depth 14 cm which was relatively rich in CuSO₄ and less concentrated in Na₂SO₄ (specific gravity, SG = 1.25 at T = 22°C). Cooling from below produced the same sequence of crystal layers, first the copper, then the sodium salt, but with the changeover occurring sooner because the composition changes were confined to a shallower fluid layer. During this process, however, the density of the lower layer was decreasing, until eventually it overturned and mixed with the upper layer. This suddenly brought the top of the crystals into contact with a Cu-rich solution, and another layer of CuSO₄·5H₂O built up on top of the previous layers, as shown in Fig. 3b. Note that though the fluid layers undoubtedly controlled the growth and sequence of the crystal layers, the thicknesses in the two states were not identical, and only indirectly related.

When a layer of mixed solution of copper and sodium sulphates, with a depth of 6 cm and composition similar to that used in the experiments already described, was cooled from above through an interface (using a metal chamber containing coolant in contact with a less dense (SG = 1.055) upper layer of cold Na₂SO₄ solution), the same sequence of crystallization was observed as in the first experiment. A compact layer of CuSO₄·5H₂O formed right across the (insulated) bottom and grew to a thickness of 4 mm. On top of this, there was a sudden change to Na₂SO₄·10H₂O crystals, which began to grow after about 18 h in a long, spikey form (rather than as a compact

mass as they did with bottom cooling) and built up to a nearly uniform thickness of 4.0 cm after about 4 days. There was no overturning in this case, because of the large density difference across the interface between the two layers.

There is no doubt that more complicated crystal sequences can be produced when more fluid layers are present initially, when there are more components which can crystallize over the temperature and composition ranges attainable in an experiment, or when the cooling conditions (for example the temperature of the base of the tank) are changed. The next case to be discussed is a three-layer system in which two solutes were present initially. The top, 12-cm deep, layer was KNO₃ solution at 24°C, SG = 1.148; the middle layer was also KNO₃ at 46°C, SG = 1.192, 4 cm deep; and the bottom layer (put in last) was CuSO₄ solution at 54°C, SG = 1.310, 4 cm deep (Fig. 1c). Cooling from the bottom (with the coolant at 8°C) produced a 5 mm thick layer of CuSO₄·5H₂O crystals in 5.5 h, during which time the three fluid layers remained distinct. Overnight, after a total of 19.5 h, the bottom interface had broken down; some of the CuSO₄·5H₂O crystals had redissolved, but a thinner layer survived, above which was an 8-mm layer of paler blue crystals, identified (by X-ray diffraction) as the double sulphate K₂SO₄·CuSO₄·6H₂O. The coolant temperature was dropped to 0°C, and after a further day (a total of 45.5 h) the second interface had broken down, and convection extended over the whole depth of the fluid. On top of the two previously observed crystal layers was a thicker layer of KNO₃ crystals, compact at the bottom but with needles on top (see Fig. 3c). Thus each time an interface broke down, the composition of the fluid in contact with the cooled bottom changed abruptly, allowing one or more new solid phases to crystallize.

An example will now be given in which the thickness of the crystal layers was more directly controlled by the depth of the compositional layers in the fluid. A 12-cm deep layer of (NH₄)₂SO₄ solution at 10°C, SG = 1.187, was set up above a 5-cm deep layer of CuSO₄ solution at 54°C, SG = 1.296 (Fig. 1d). The interface was initially sharp, but it was stirred gently to produce a transition region 4 cm thick. The coolant circulating through the base was brought down from room temperature to 5°C; on this occasion the thin insulating layer (usually built into the base of the tank to maintain a uniform temperature over the whole area) was left out, and the coldest temperature were therefore close to the back face. After 1.5 h, a layer of

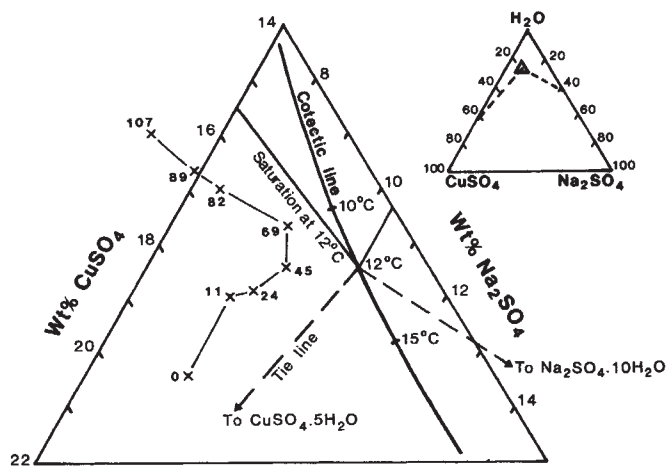


Fig. 2 Part of the phase diagram of the CuSO₄-Na₂SO₄-H₂O system, showing the cotectic line (representing solutions in equilibrium with both CuSO₄·5H₂O and Na₂SO₄·10H₂O), and the saturation lines for the two components separately. The region detailed is contained in the small triangle on the full phase diagram sketched at the upper right. The evolution of the bulk composition of the fluid, during the experiment corresponding to Figs 1a and 3a, is also plotted. Times (in hours since the start of cooling) are marked against the measured points.

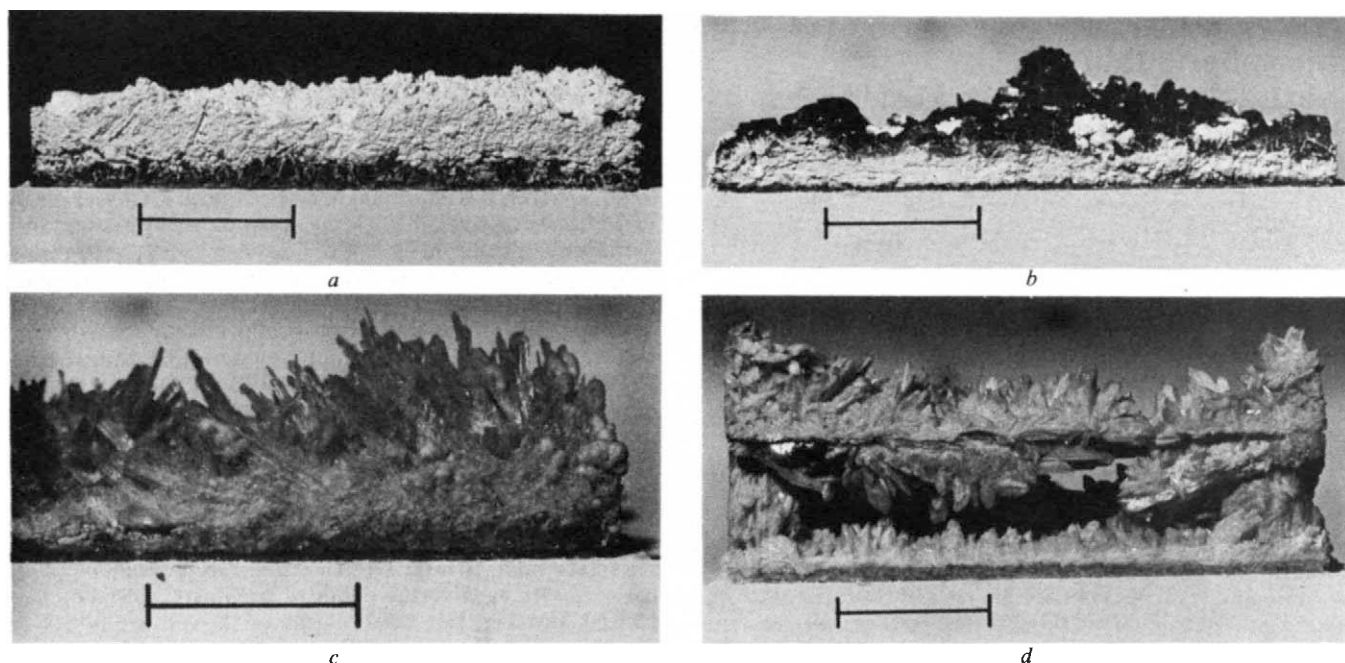


Fig. 3 *a*, Horizontal layers of first $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, then $\text{Na}_2\text{SO}_4 \cdot 10\text{H}_2\text{O}$ crystals, produced by cooling from below, as the composition of a layer of mixed solution of Cu and Na sulphates evolved with time (see Figs 1*a* and 2). The photograph was taken after the sodium sulphate crystals had dehydrated. (All scale bars, 5 cm.) *b*, The crystal structure produced by cooling a two-layer fluid system from below. The lower layer consisted of a mixture of Cu and Na sulphates similar to that recorded in Fig. 1*a*, and the first two crystal layers grew as described in *a*. The density of the lower layer decreased until overturning occurred, and another layer of $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ then grew on top from the Cu-rich solution newly in contact with the crystals. The first indication of a second layer of sodium sulphate is also seen above this. *c*, The crystal structure produced by cooling a three-layer fluid system from below. The upper two layers were KNO_3 (with different concentrations and temperatures, as detailed in Fig. 1*c*), and the lowest was hot CuSO_4 solution. The crystal layer which first formed over the bottom was $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$. As the two interfaces broke down successively, and the composition of the fluid in contact with the top of the crystals changed abruptly, layers of the double sulphate crystals ($\text{K}_2\text{SO}_4 \cdot \text{CuSO}_4 \cdot 6\text{H}_2\text{O}$) and then KNO_3 crystals formed on top. *d*, Crystal layers which correspond directly to the pre-existing compositional layering in the fluid from which they grew (Fig. 1*d*). A 5-cm deep layer of hot CuSO_4 solution was placed below a 12-cm deep layer of less dense cool $(\text{NH}_4)_2\text{SO}_4$ solution, and the interface thickened by stirring. Cooling from below produced a uniform layer of $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ on the bottom and a horizontal band of the double sulphate $(\text{NH}_4)_2\text{SO}_4 \cdot \text{CuSO}_4 \cdot 6\text{H}_2\text{O}$, with a small fraction of $(\text{NH}_4)_2\text{SO}_4$, in the interfacial region. When overturning occurred, to produce a single convecting fluid layer, further growth of the double salt took place, extending away from the previous deposit.

$\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ crystals had begun to grow over the whole floor, and a few paler blue crystals (probably the double sulphate $(\text{NH}_4)_2\text{SO}_4 \cdot \text{CuSO}_4 \cdot 6\text{H}_2\text{O}$) were observed on the walls and corners at the centre of the transition layer. No observations were made overnight, but by 19 h the fluid in the tank had overturned, and the structures shown in Fig. 3*d* were observed (looking from the back—above the most strongly cooled region). There is a layer of copper sulphate on the base, with above it a uniform layer of the double sulphate crystals, mixed with a small fraction of $(\text{NH}_4)_2\text{SO}_4$ crystals. A horizontal layer of the double sulphate plus some ammonium sulphate is also visible at the level of the original interface, and it can be deduced that this is where the composition of the mixed solution initially favoured the formation of these two phases; further crystallization has occurred upwards and downwards from this level following the overturn and complete mixing.

Thus at least three mechanisms have been identified whereby a series of horizontal crystal layers of nearly constant thickness and abruptly changing composition can be produced in a layered multicomponent fluid system cooled from below. First, the composition of a single isolated layer may evolve (more rapidly than would otherwise be the case in a very deep well-mixed container), so that crystallization of a new phase with different kinetics becomes possible above the crystals originally formed on the bottom. Second, the decrease in density of the residual fluid which accompanies crystallization can lead to a sudden overturning and mixing, which rapidly changes the composition of the deeper combined layer in contact with the top of the crystals previously formed. Third, the interfacial region between fluid layers of different composition can provide an environment

in which some species will crystallize preferentially, at higher temperatures than any crystals can form in either layer separately. In general the thickness of the various crystal layers is determined only indirectly by the depth of the fluid layers, but in the third case there is a direct relationship between the two scales.

The dynamical principles sketched here should be directly applicable to crystallizing magmas. The justification for transferring results from one fluid system to another, and the associated scaling problems, have been discussed elsewhere^{1,5} and the need for this type of study has already been recognized in the recent literature^{9,10} on layered intrusions. Further work in progress is aimed at providing a more quantitative understanding of the evolution of crystal layers and the solutions in which they are growing, so that the laboratory results can be generalized and used predictively for more complex magmatic systems.

We thank D. L. Corrigan and R. Wylde-Browne for assistance at all stages of these experiments, and P. W. Berrie for help with the X-ray diffraction analyses.

Received 11 March; accepted 15 June 1982.

1. Chen, C. F. & Turner, J. S. *J. geophys. Res.* **85**, 2573–2593 (1980).
2. Turner, J. S. *Nature* **285**, 213–215 (1980).
3. McBirney, A. R. *J. Volcanol. geotherm. Res.* **7**, 357–371 (1980).
4. Huppert, H. E. & Turner, J. S. *Earth planet. Sci. Lett.* **54**, 144–152 (1981).
5. Turner, J. S. & Gustafson, L. B. *J. Volcan. geotherm. Res.* **11**, 93–125 (1981).
6. Huppert, H. E. & Sparks, R. S. *J. Contr. Miner. Petrol.* **75**, 279–289 (1980).
7. Huppert, H. E., Turner, J. S. & Sparks, R. S. *J. Earth planet. Sci. Lett.* **57**, 345–357 (1982).
8. Wager, L. R. & Brown, G. M. *Layered Igneous Rocks* (Oliver and Boyd, Edinburgh, 1967).
9. McBirney, A. R. & Noyes, R. M. *J. Petrol.* **20**, 487–554 (1979).
10. Irvine, T. N. in *Physics of Magmatic Processes* (ed. Hargreaves, R. B.) Ch. 8 (Princeton University Press, 1980).

Lithospheric thinning associated with rifting in East Africa

Richard F. Wendlandt & Paul Morgan

Lunar and Planetary Institute, 3303 NASA Road One, Houston, Texas 77058, USA

A fundamental mechanism of some continental rifting processes appears to be the transmission of thermal energy into the lithosphere by asthenospheric upwelling (the 'active' mechanism of Sengör and Burke¹), for which the East African rift system is a classic example. The abundance and variety of magma types of predominantly alkaline affinities in the East Rift which derive from the thermal perturbation suggest that studies of igneous petrogenesis may offer a method of investigating the physical as well as the chemical conditions of the upwelling. Systematic variations in the temporal-spatial-compositional relations of igneous rocks associated with rifts are becoming well documented for several different rifts²⁻¹³, and it is generally observed that the degree of silica-saturation of lavas within these rifts increases with time, while incompatible element contents decrease. Furthermore, at any given instant, lavas erupted within the rifts are less silica-undersaturated than lavas erupted outside the rift. These observations are consistent with decreasing depths of origin of magmas with time as would be predicted for magmagenesis associated with an upwelling source region. Here, by deducing the physical conditions of origin of various magmatic products associated with a thermotectonic event in East Africa, the rate of upwelling is calculated and compared with the results of a simple thermal model of lithospheric thinning.

The diversity of eruptive rock types in the Kenya Dome region of the East African rift, corresponding to depths of origin ranging from >170 km to the lower crust (25–30 km), combined with an abundance of dated occurrences permits calculation of the rate of ascent of the thermal perturbation (Fig. 1). As a first approximation it is assumed that the magmas being considered in the analysis were generated and/or segregated from the lithosphere–asthenosphere boundary. The points in Fig. 1 were selected so as to minimize the depth of origin of magmas for a given age. The physical and chemical conditions of origin of the more primitive alkalic varieties of magmas and the flood volcanism in the East Rift are emphasized inasmuch as it is believed that the compositions of these rocks can be related to the depth from which they erupted. The deepest and oldest point in Fig. 1 is defined by the most recent diamond-bearing kimberlite eruption in northern Tanzania, 41 Myr (ref. 14); the depth of origin is based on the intersection of the theoretical and experimentally derived conditions for kimberlite genesis with the diamond–graphite equilibria¹⁵⁻¹⁸, and represents a minimum estimate. The next younger point in Fig. 1, 31 Myr, is based on the oldest known carbonatite associated with rifting, Tororo in eastern Uganda¹⁹; the range of depths of origin is based on partial melting experiments on carbonated, and carbonated and hydrated peridotite^{16,20-27}. The point at ~22 Myr represents the initiation of eruption of ~50,000 km³ of Miocene flood basalts in northwestern and western Kenya⁵; the depth of origin is estimated from experimental results²⁸⁻³⁰. The two youngest points represent the onset of eruption of flood phonolites and trachytes, respectively, in the East Rift in southwestern Kenya and northern Tanzania³¹; the depths of origin are constrained by experiments in model peridotite systems^{16,32} and thermobarometric estimates³³. Approximately 65,000 km³ of flood phonolite and trachyte were erupted^{5,34}. The depth of origin estimates for the flood trachytes are probably realistic, being situated in the lower crust, although any model assumed in deriving the depth estimate will be, at best, contentious: of greater importance to the analysis is the observation that

significant ascent of mantle isotherms must have occurred to enable the accumulation of 10⁴–10⁵ km³ of melts of trachytic composition.

The magmatic events forming the basis of the analysis comprise >10⁵ km³ of eruptive rock, an amount approaching estimates of the total volume associated with the Kenya Dome. The data in Fig. 1 are represented by points with respect to time; however, it is important to emphasize, at least with regard to the flood eruptives, that it is the initiation of magmatic events, perhaps comprising many tens of flows and lasting up to several million years, which forms the basis for the analysis. While the general trend of melt composition variation with time is well documented⁴⁻⁹, the volcanic stratigraphy is complex with intermingling of eruptives of different compositions, and, thus, of different depths of origin, particularly for younger magmas erupted from shallower depths; that is, the progression of deep sourced eruptives to those from shallower depths is not without exceptions, perhaps attesting to complex convective processes in the asthenosphere (textural and structural evidence of episodic deformation and melting that is believed to have originated in an diapirically rising mass has been reported for ultramafic nodules from Lashaine Volcano, Tanzania)³⁵. If all the available ages versus depth of origin for volcanics were plotted, however, they would cluster on the concave side of the envelope (Fig. 1), without affecting the exponential trend of the minimum depth of origin for a given age. A systematic decrease in the rate of ascent of asthenosphere is evident (Fig. 1). Complete thinning of the lithospheric upper mantle by asthenospheric upwelling is accomplished over a time interval of ~30–40 Myr.

To understand better the lithospheric attenuation predicted by the analysis of eruptives in East Africa, a simple thermal model of lithospheric thinning has been developed. The lithosphere–asthenosphere boundary (LAB) is assumed to lie at or above the top of the upper thermal boundary layer of mantle convection, with its temperature at the mantle solidus. Heat

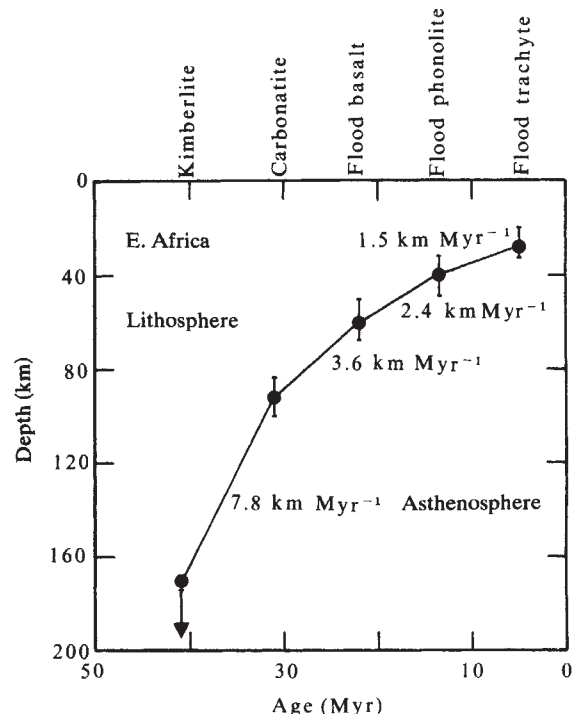


Fig. 1 Asthenosphere ascent rate associated with active type rifting in the Kenya Dome portion of the East African rift system as estimated by changes in the depth of origin of magmatic products with time. Data points are discussed in the text. The ranges of experimentally determined depths of origin are shown by brackets; the selected points are considered to be the best estimates except for the depth of origin of kimberlite which is a minimum estimate.

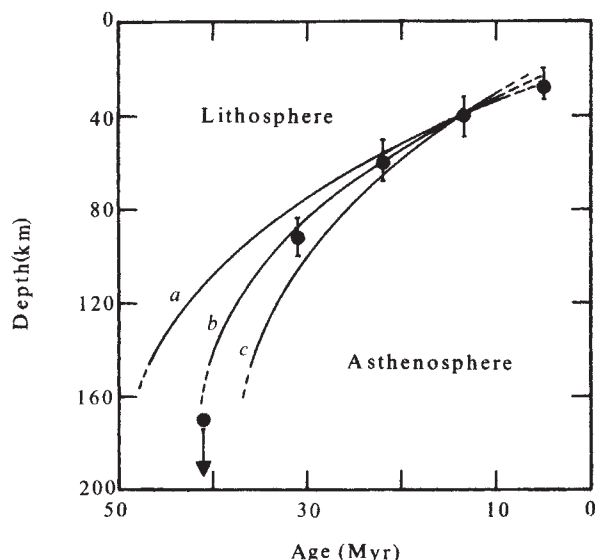


Fig. 2 Asthenosphere ascent, constrained to pass through the phonolite (40 km, 13.5 Myr) data point, calculated from thermal thinning model for 50 km radius lithosphere-asthenosphere boundary perturbation and different temperature gradient increases across the lithosphere-asthenosphere boundary: a, $70^{\circ}\text{C km}^{-1}$; b, $60^{\circ}\text{C km}^{-1}$; c, $50^{\circ}\text{C km}^{-1}$. Curves for 100 km radius perturbation are similar for 65, 55 and $45^{\circ}\text{C km}^{-1}$ gradient increases respectively. Data points from Fig. 1 are shown for comparison.

transfer in the lithosphere is assumed to be by conduction only. In stable conditions, heat is conducted through the lithosphere at the same rate that it is input through the boundary layer. If locally the heat supply is increased (mantle 'hotspot'), however, temperatures will increase in the lithosphere, and a perturbation in the LAB will ascend to maintain its position on the mantle solidus. It is this ascent rate, as a function of the undisturbed lithosphere geotherm, the mantle solidus, and the increased heat supply, that we seek to determine.

By simplifying the geometry of the LAB perturbation to be spherical over the heat source, the following equation has been derived for the relationship between the increased heat supply (increased gradient dT'/dr), the temperature increase (T') at the LAB, and the ascent rate (W) of the LAB perturbation, from the solution for the temperature from a stationary constant-strength point-source in a medium moving at constant velocity³⁶:

$$\frac{dT'/dr}{T'} = \left[\frac{(1 + Wr/k) \exp(-Wr/k)}{r^2} + \frac{(1 + W(r+2z)/k) \exp(-W(r+2z)/k)}{(r+2z)^2} \right] / \left[\frac{\exp(-Wr/k)}{r} - \frac{\exp(-W(r+2z)/k)}{r+2z} \right] \quad (1)$$

where r is the radius of the LAB perturbation, z is the depth of its shallowest point and k is the thermal diffusivity of the lithosphere. Although this equation is derived for a constant velocity ascent, a parameter study has shown that it can be used to determine instantaneous velocities at various depths, even if the ascent rate is variable, for radii ≥ 40 km and ascent velocities $\geq 1 \text{ km Myr}^{-1}$ ($3.17 \times 10^{-11} \text{ m s}^{-1}$), for a typical mantle diffusivity of $1 \text{ mm}^2 \text{ s}^{-1}$. This approximation is satisfactory because at velocities $> 1 \text{ km Myr}^{-1}$ the ascent rate is effectively faster than the thermal conduction rate above the LAB.

The temperature increase in equation (1) is the temperature required to raise the lithosphere geotherm to the solidus. This has been calculated from a shield geotherm, considered to be

appropriate for East Africa before rifting, given by model C1 of Sclater *et al.*³⁷, and a mantle solidus derived from published experimental data¹⁶. This mantle solidus was selected because it is internally consistent over a large pressure range, and is a model for the genesis of alkalic magmas, carbonatites and kimberlites, assuming the presence of small amounts of H_2O and CO_2 in the mantle. A thermal diffusivity of $1 \text{ mm}^2 \text{ s}^{-1}$ was assumed, and equation (1) was solved iteratively for ascent velocity at 10-km depth intervals from 150 to 30 km, assuming different radii for the LAB perturbation and different temperature gradient increases across the LAB.

Excellent fits to the experimentally derived LAB ascent rates were obtained for radii of 50 and 100 km and constant gradient increases within the range $55\text{--}60^{\circ}\text{C km}^{-1}$ over the LAB ascent depth range, as shown in Fig. 2. These radii can reconcile the lateral extent of the eruptives in East Africa³⁸, as the ascent rate is not sensitive to the radius if the latter is > 100 km. The range of gradient increase is more difficult to evaluate, but is not unrealistic, being of the same order in a fixed heat source-lithosphere reference frame as that required in a moving reference frame to reheat the Pacific plate to a bathymetric age of 25 Myr over hotspots, as reported by Detrick and Crough³⁹. The ascent rate is sensitive to changes in dT'/dr ; a $1^{\circ}\text{C km}^{-1}$ change in gradient approximately results in a 0.55 Myr change ($\sim 1.5\%$) in the total ascent time from 150 to 30 km. The apparent 'exponential' decrease in ascent rate with time is predicted even with a constant heat supply (constant dT'/dr). This decrease is due to the increasing differential (T') between the geotherm and the solidus as the depth decreases, and the increasing rate of heat loss to the surface as the thermal wave in front of the LAB perturbation impinges on the surface.

Two direct predictions result from the thermal model. The rapid ascent rate of the thermal perturbation indicates little thermal disturbance above the LAB. The experimental data indicate that the perturbation reached a depth of 30 km 4.8 Myr ago, hence only a small surface heat flow anomaly is predicted outside the areas of Quaternary volcanism and plutonism. This prediction is consistent with surface heat flow data across the Kenya rift⁴⁰. The thermal perturbation should cause thermal expansion in the lithosphere-asthenosphere column as it rises, however, resulting in uplift contemporaneous with lithospheric thinning. Using a volume thermal expansion coefficient of $3 \times 10^{-5} \text{ }^{\circ}\text{C}^{-1}$ the model predicts a mean decrease in density of 0.035 g cm^{-3} from 160 to 20 km depth during thinning, which, assuming the lithosphere to be supported buoyantly by the asthenosphere⁴¹, results in an uplift of a little over 1.5 km. This uplift is in good agreement with the uplift determined from geological criteria in Kenya⁵.

The temporal-spatial-compositional relations of eruptives in the Kenya Dome region of the East African rift system suggest a shallowing of magma source regions with time, an observation implying encroachment of a thermal anomaly as might be expected from asthenospheric upwelling. Thermal modelling indicates that the rate of this upwelling, and the experimentally determined decrease in this rate at shallower depths can be explained by a constant supply of heat from the asthenosphere, with a magnitude similar to that deduced for other inferred mantle hotspots. The experimentally determined and theoretically modelled lithospheric thinning gives results consistent with observed surface heat flow and uplift across the Kenya rift, and supports the concept of a convective upwelling of the asthenosphere driving rifting in East Africa.

We thank R. J. Phillips, W. Harrison and G. Wadge for helpful discussions and constructive criticism. This work was carried out at the Lunar and Planetary Institute, operated by the Universities Space Research Association under contract NASW 3389 from NASA. Lunar and Planetary Contribution No. 478.

Received 22 February; accepted 8 June 1982.

1. Sengör, A. M. C. & Burke, K. *Geophys. Res. Lett.* **5**, 419-421 (1978).
2. Gass, I. G. *Phil. Trans. R. Soc. A* **267**, 369-381 (1970).
3. Mohr, P. A. *Bull. Volcanol.* **34**, 141-157 (1970).

4. Baker, B. H. *et al.* *Tectonophysics* **11**, 191–215 (1971).
5. Baker, B. H. *et al.* *Geol. Soc. Am. Spec. Pap.* **136**, (1972).
6. King, B. C. & Chapman, G. R. *Phil. Trans. R. Soc. A* **271**, 185–208 (1972).
7. Williams, L. A. J. *Tectonophysics* **15**, 83–96 (1972).
8. Lippard, S. J. & Truckle, P. H. in *Petrology and Geochemistry of Continental Rifts* (eds Neumann, E.-R. & Ramberg, I. B.) 123–131 (Reidel, Dordrecht, 1978).
9. Norry, M. J. *et al.* *Phil. Trans. R. Soc. A* **297**, 259–271 (1980).
10. Lipman, P. W. *Bull. geol. Soc. Am.* **80**, 1343–1354 (1969).
11. Dungen, M. A. *et al.* in *Pap. Conf. on the Processes of Planetary Rifting*, 137–140 (Lunar and Planetary Institute, Houston, 1981).
12. Duncan, R. A. *et al.* *Nature* **239**, 82–86 (1972).
13. Brooks, C. K. & Rucklidge, J. C. *Lithos* **7**, 239–248 (1974).
14. Davis, G. L. *Yb. Carnegie Instn Wash.* **77**, 895–897 (1978).
15. Eggler, D. H. & Wendlandt, R. F. in *Kimberlites, Diatremes, and Diamonds: Their Geology, Petrology, and Geochemistry* (eds Boyd, F. R. & Meyer, H. O. A.) 330–338 (American Geophysical Union, Washington DC, 1979).
16. Wendlandt, R. F. & Eggler, D. H. *Am. J. Sci.* **280**, 421–458 (1980).
17. Wyllie, P. J. *J. geophys. Res.* **85**, 6902–6910 (1980).
18. Kennedy, C. S. & Kennedy, G. C. *J. geophys. Res.* **81**, 2467–2470 (1976).
19. King, B. C. *et al.* *J. geol. Soc.* **128**, 173–205 (1972).
20. Wyllie, P. J. & Huang, W. L. *Nature* **257**, 297–299 (1975).
21. Wyllie, P. J. & Huang, W. L. *Geology* **3**, 621–624 (1975).
22. Wyllie, P. J. & Huang, W. L. *Contr. Miner. Petrol.* **54**, 79–107 (1976).
23. Eggler, D. H. *Yb. Carnegie Instn. Wash.* **74**, 468–474 (1975).
24. Eggler, D. H. *Geology* **4**, 69–72 (1976).
25. Eggler, D. H. *Am. J. Sci.* **278**, 305–343 (1978).
26. Wendlandt, R. F. & Mysen, B. O. *Am. Miner.* **65**, 37–44 (1980).
27. Wendlandt, R. F. & Eggler, D. H. *Am. J. Sci.* **280**, 385–420 (1980).
28. Green, D. H. *Phys. Earth planet. Inter.* **3**, 221–235 (1970).
29. Thompson, R. N. *Contr. Miner. Petrol.* **45**, 317–341 (1974).
30. Jaques, A. L. & Green, D. H. *Contr. Miner. Petrol.* **73**, 287–310 (1980).
31. King, B. C. & Chapman, G. R. *Phil. Trans. R. Soc. A* **271**, 185–208 (1972).
32. Kushiro, I. *J. geophys. Res.* **73**, 619–634 (1968).
33. Irving, A. J. & Price, R. C. *Geochim. cosmochim. Acta* **45**, 1309–1320 (1981).
34. Lippard, S. J. *Lithos* **6**, 217–234 (1973).
35. Pike, J. E. N. *et al.* *J. Geol.* **88**, 343–352 (1980).
36. Carslaw, H. S. & Jaeger, J. C. *Heat Conduction in Solids*, 266–267 (Oxford University Press, London, 1959).
37. Sclater, J. G. *et al.* *Rev. Geophys. Space Phys.* **18**, 269–311 (1980).
38. Wendlandt, R. F. in *Pap. Conf. on the Processes of Planetary Rifting*, 126–133 (Lunar and Planetary Institute, Houston, 1981).
39. Detrick, R. S. & Crough, S. T. *J. geophys. Res.* **83**, 1236–1244 (1978).
40. Morgan, P. & Wheildon, J. *J. geophys. Res.* (submitted).
41. Crough, S. T. & Thompson, G. A. *J. geophys. Res.* **81**, 4857–4862 (1976).

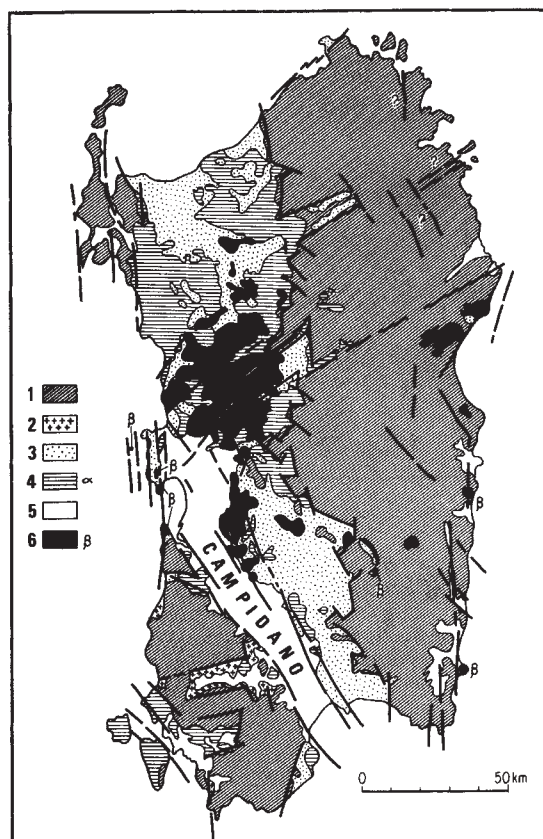


Fig. 1 Geological map of Sardinia. 1, Basement with its Mesozoic and Eocene cover; 2, Cixerri Formation; 3, Oligocene-Miocene rifted basin; 4, Oligocene-Miocene calc-alkaline volcanics; 5, continental Pliocene-Quaternary basin; 6, alkaline Pliocene-Quaternary volcanics.

Oligo-Miocene rift of Sardinia and the early history of the Western Mediterranean Basin

A. Cherchi

Istituto di Geologia, Paleontologia e Geografia Fisica,
Università di Cagliari, via Trentino 51, 09100 Cagliari, Italy

L. Montadert

Istituto di Geologia, Paleontologia e Geografia Fisica,
Università di Cagliari, via Trentino 51, 09100 Cagliari, Italy

The geodynamic evolution of the Western Mediterranean Basin, in spite of many studies, is still uncertain. There is some consensus for interpreting this basin as a kind of small oceanic marginal basin. Its opening has generally been related to a subduction process which was active during the Oligocene-Miocene somewhere east of Sardinia-Corsica¹⁻⁷. As the margins of the basin are deeply buried below Miocene-to-present sediments, direct lithological and stratigraphical data which could explain the events responsible for its formation are rare⁸⁻¹⁰ or missing altogether. To obtain such data, detailed field studies have been undertaken in Sardinia (Fig. 1), and the first results are presented here. This approach is justified by the fact that in that island, Oligocene and Miocene sediments were deposited in a rift (*fossa tettonica sarda* of Verdabasso¹¹), which is the easternmost arm of the complex rift system that affected the European plate during Oligocene and Miocene times. One of these arms evolved towards a small oceanic basin—the Western Mediterranean or Algero-Provençal Basin—while others such as the Gulf of Valencia and the Sardinia rift aborted and remained at the rift stage. Its exceptional exposures make it possible to examine the Sardinia rift to clarify the sequence of events which created it, and to establish a sedimentological model which we believe is directly applicable to the Western Mediterranean Basin.

The Sardinian Oligocene-Miocene rift, which crosses all of Sardinia's 220 km on a N-S trend, must be distinguished from the Pliocene-Quaternary graben of Campidano; the two are often confused. Figures 1, 2a and 3 show that the Campidano graben is superimposed on the southwestern part of the Oligocene-Miocene graben, but in an oblique NW-SE direction. It corresponds to an entirely different event, which affected the Central and Western Mediterranean¹². Detailed structural and sedimentological observations, supported by a good biostratigraphy, can therefore be made of the eastern part of the Oligocene-Miocene rift which appears as a typical rift edge. The Palaeozoic basement of Sardinia, sometimes including its Mesozoic and Eocene cover, is downfaulted by a succession of tilted blocks. The vertical lowering of the pre-rift basement is estimated at at least 1,200 m, and probably much more. The complicated fault pattern can be explained by the complex structure and lithology of the basement. While the general trend of the rift is N-S, local faults vary from N-S to NNW-SSE, and are interrupted by transverse accidents. Among the succession of faults which downfaulted the basement, a master fault is generally observed whose throw can be as wide as 500 m (Fig. 2b). It is clearly evident in the morphology of the Sardinian rift, and explains the wall aspect on the eastern edge. Tilted blocks with shorter throw are observed westward and eastward of this master fault. The tectonic style is identical to those described in various intracontinental rifts and on some rifted passive continental margins¹³. This complex pattern controlled the palaeogeography of the basin and the nature and deposition of the sediments along the rift. The width between the two master faults delineating the main trough is 40–50 km.

The Sardinian rift has been infilled by Oligocene-Miocene continental-to-marine sediments ~1 km thick. Volcanics occurring from the late Oligocene, 29.9 ± 1.5 Myr to the middle

Fig. 2 *a*, Schematic geological cross-section from Menorca to Sardinia through the Western Mediterranean basin. Modified from Bijou-Duval *et al.*⁵. The section crosses western Sardinia near Oristano. *b*, Geological cross-section on the eastern edge of the Oligocene-Miocene rift showing the tilted block of Isili bounded westward by a master fault. Conglomerates, sandstones and limestones are Oligocene-Aquitainian continental and marine syn-rift deposits. The black dotted line corresponds to the early Burdigalian post-rift marls.

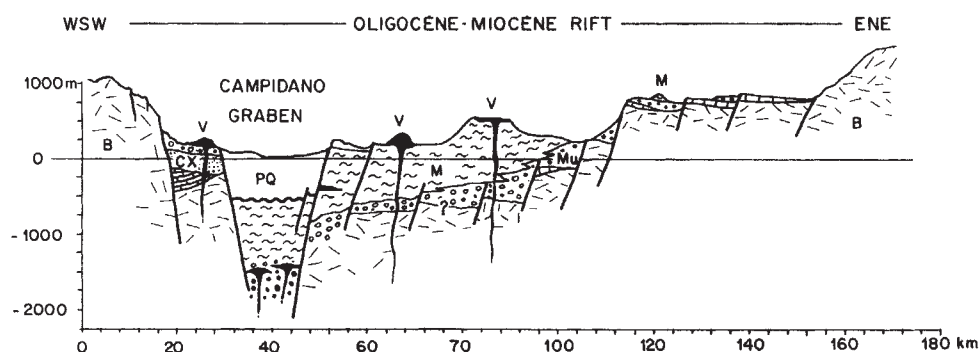
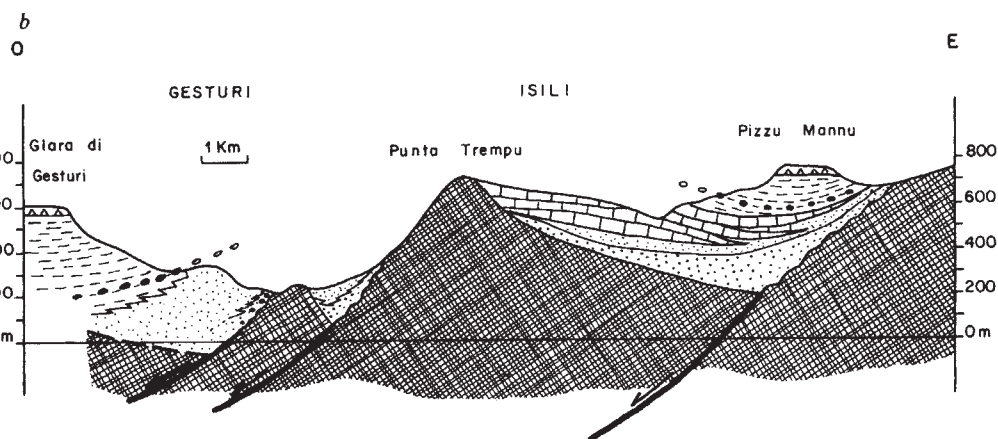
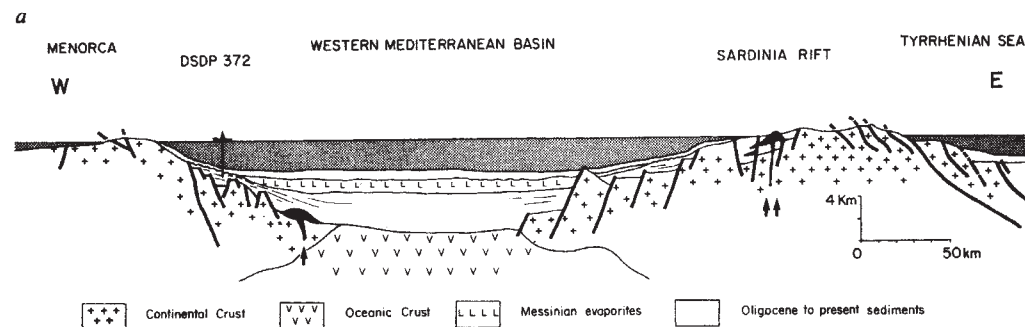


Fig. 3 Schematic geological cross-section through the Oligocene-Miocene rift of Sardinia, with the superimposed Pliocene-Quaternary Campidano graben. B, Palaeozoic basement with sometimes its Mesozoic and Eocene cover. Cx, Cixerri Formation; Mu, Oligocene-Miocene synrift Ussana Formation; M, late Oligocene-Miocene marls; PQ, Pliocene-Quaternary; V, volcanics.

Miocene, 13–14 Myr (refs 6–14) can be up to 1 km thick. They are believed to be the products of the volcanic arc linked with the subduction process which was active east of Sardinia. The sediments can be studied in their structural context to determine the timing of the formation, the evolution of the Sardinian rift and the age and nature of the prerift, synrift and postrift deposits.

Before the formation of the Western Mediterranean Basin, the Corsica Sardinia block joined the European plate. Lithological and biostratigraphical data demonstrate this proximity until the early to middle Oligocene. Indeed the continental Cixerri Formation (~300 m thick) outcrops in southwestern Sardinia where it is unconformable to various strata: Palaeozoic, Mesozoic and mainly Lower-base-Middle Eocene. This unconformity corresponds to the Eocene Pyrenean compressional phase recognized and recorded in Sardinia^{15,16} with a NW-SE shortening direction¹⁷. The microfaunal composition and the direction of currents measured on these clastics¹⁸, demonstrate a provenance from northeastern Iberia^{19–21} for the material constituting the Cixerri Formation. These observations, and the absence of pebbles of local origin, suggest that the clastics were deposited after the Pyrenean phase and before the initiation

of the rift system of the Western Mediterranean. The Cixerri Formation is dated Lutetian at its base by Charophyta and pollen assemblage²². It is cut by andesites, the older ones being dated 29.9 ± 1.5 Myr. It can be compared with conglomerates in Provence and in Menorca¹⁹.

The fundamental structural change corresponding to the formation of the rift is recorded by the onset of deposition of the thick clastic Ussana Formation²³. Its age is post 29.9 ± 1.5 Myr, early Middle Oligocene, in accordance with a Stampian age

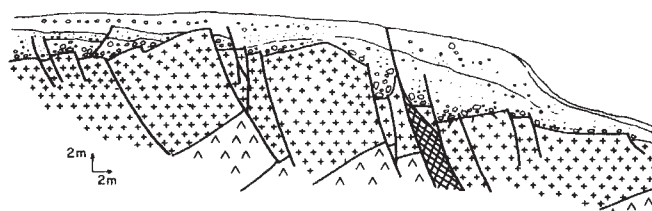


Fig. 4 Detail of the rift tectonics with faulting synchronous of continental conglomerates deposition (north of Cagliari).

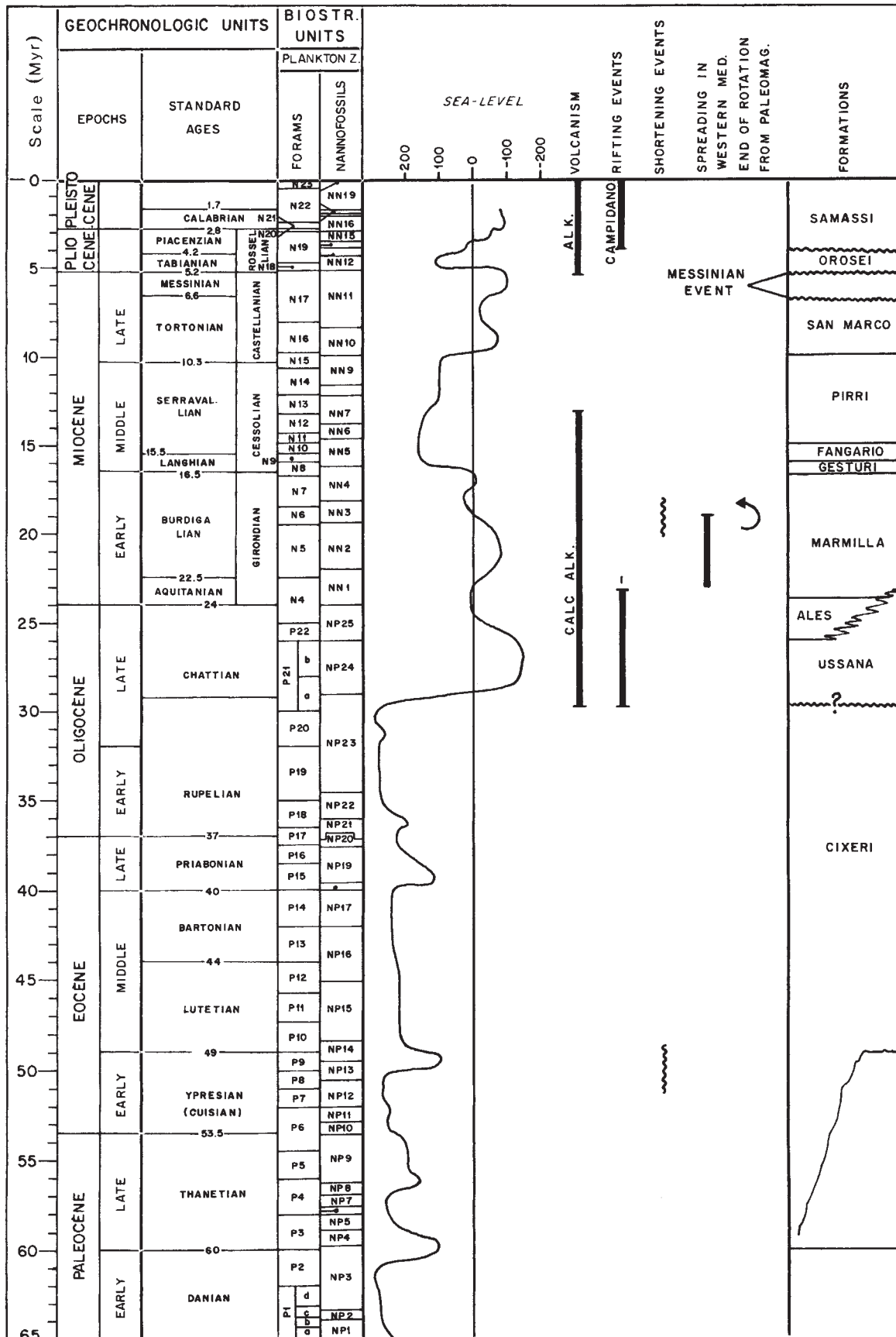


Fig. 5 Main geological events in Sardinia. Time scale and eustatic changes of sea level from ref. 38.

established for grabens in Languedoc (Southern France)²⁴. The creation of new reliefs and fault escarpments caused intensive erosion.

First, continental breccias and heterometric conglomerates (from a few centimetres to several metres) embedded in a

red-violine sandy-clayey matrix were deposited on and along these escarpments. Their structure is typical of slope deposits, alluvial fans or torrential deposits and they are clearly contemporaneous with faulting (Fig. 4). These observations demonstrate that in its early phase, rifting occurred in a continental

environment. Then, fluvial conglomerates with sandy matrices sometimes including breccias and intercalations of marine limestones with corals, oysters, bryozoans, red algae and pectinids, indicate progressive invasion of the rift by the sea. Thick sandstones are also present. This Ussana Formation, as much as 500 m thick, developed along the edge of the rift, its complex distribution of deposits controlled by the nature and topography of the basement. Locally, along the crest of some tilted blocks, shallow water carbonates made of red algae and >200 m thick were deposited (Fig. 2b). As it approaches the axis of the rift the upper Ussana Formation rapidly changes to open marine marls rich in pelagic microfossils (Ales and Marmilla Formations)²⁵, while on the flank, shallow-water limestones are present. The change from the continental, fluvial or shallow marine environment to the deeper basin is indicated by intercalations of sands, conglomerates and breccias, with blocks up to several metres, deposited as turbidites and debris flows.

Rifting originated in a continental environment during the early Middle Oligocene. But these data, the pelagic microfossils present in the oldest open marine marls (which are Chattian—P22–NP25—in age), and the shallow-water limestones on the flanks (dated with *Lepidocyclina* and *Miogyopsina* gr. *bantamensis*) indicate that the sea invaded the rapidly subsiding axis of the rift during the late Oligocene. The deposition of the Ussana Formation continued until the Aquitanian (N4). Quantitative study²⁶ of benthic and planktonic foraminifera has resulted in some estimates of the bathymetry at that time. For the Aquitanian deposited at the foot of the master fault, palaeodepths of ~1–1.3 km have been determined, while palaeodepths of 200–300 m were found in Aquitanian deposited higher up. Thus the sea bottom at that time was characterized by high relief.

Deposits equivalent to the Ussana, Ales and Marmilla Formations have been found around the Western Mediterranean. For example, conglomerates and late marine Oligocene have been found in exploration wells in the Camargue trough²⁷ and in the Gulf of Lions¹⁰ and by submarine sampling on the continental slope of Provence^{28–30}. In the Balearic Islands synrift deposits of early Miocene age have also been described^{31,32}. Since the Aquitanian, Miocene deposits are essentially marine marls throughout the basin, with the exception of a sandy and carbonate episode in the Upper Serravallian–Lower Tortonian. Calc-alkaline volcanics^{33–35} are interbedded throughout the Oligocene–Miocene deposits, and are dated from 29.9±1.5 Myr to 13–14 Myr (refs 6, 14). They were emplaced either in the air (ignimbrites) or under the sea (pillow-lavas). Associated with them, thick and frequent series of tuffites were deposited. Major events are shown in Fig. 5. Late Miocene-to-present deposits are not considered here.

Thus there is a good set of evidence to support the thesis that rifting in Sardinia, and elsewhere in the European plate, began in the middle Oligocene around 30 Myr ago in a continental environment. At the end of the Oligocene, the trough was deep enough to allow invasion by the sea. The precise estimate of the end of rifting in the Aquitanian 23–24 Myr ago, is based on the following evidence: (1) the typical arrangement of Aquitanian sediments in half grabens, demonstrating that they are synrift deposits; (2) the vertical offset of several hundred metres of shallow marine lower Aquitanian limestones along the master fault; (3) the sealing of faulted blocks by Aquitanian sediments. Further extension with accretion of new oceanic crust was confined to the Western Mediterranean basin. There is no sedimentological evidence in the Miocene marine marls of Sardinia to indicate the end of oceanic accretion in the Western Mediterranean marginal basin, that is the end of rotation of the Corsica–Sardinia microplate. But an intra-Burdigalian compressional phase was recently determined in Sardinia. The shortening direction trends NE–SW¹⁷, as in adjacent areas of Italy and Tunisia. We interpret this as a change in the stress field in the area at that time. The tensional regime in the Western Mediterranean ceased with the end of spreading

and rotation of Corsica–Sardinia, because collision occurred east of Sardinia. This timing is in good accordance: (1) with the reinterpretation of the K/Ar ages and palaeomagnetic data of the volcanics of Sardinia³⁶, as a 19-Myr age (intra-Burdigalian) is proposed for the end of rotation; and (2) with the 19-Myr age of the Tristanite dredged from the axis of the Ligurian sea³⁷.

We thank Drs De Graciansky, Chenet, Lemoine, Heydman and Grosjean for useful discussions.

Received 24 March; accepted 23 June 1982.

1. Boccaletti, M. & Guazzone, G. *Mem. Soc. Geol. It.* **11**, 201–216 (1972).
2. Boccaletti, M. & Guazzone, G. in *Paleogeografia del Terziario sardo nell'ambito del Mediterraneo occidentale* (ed. Chérchi, A.) 57–68 (1974).
3. Auzende, J. M., Bonnin, J. & Olivet, J. L. *J. geol. soc. Lond.* **129**, 607–620 (1973).
4. Biju-Duval, B. et al. in *The Geology of Continental Margins* (eds Burk, C. A. & Drake, C. L.) 695–721 (1974).
5. Biju-Duval, B., Letouzey, J. & Montadert, L. *Init. Rep. DSDP Leg 42A*, 951–984 (1978).
6. Bellon, H., Coulon, C. & Edel, J. B. *Bull. Soc. géol. Fr.* **19**, 825–831 (1977).
7. Tapponnier, P. *Bull. Soc. géol. Fr.* **19**, 437–460 (1977).
8. Ryan, W. B. F. et al. *Init. Rep. DSDP Leg 13*, 1–1448 (1973).
9. Hsu, K. J. et al. *Init. Rep. DSDP Leg. 42A*, 1 (1978).
10. Cravatte, J. et al. *Not. Mém. Compagnie Fr. Pétrole* **11**, 209–274 (1974).
11. Vardabasso, S. *Geol. Rdsch.* **53**, 613–630 (1963).
12. Mauffret, A. et al. *Bull. Am. Ass. petrol. Geol.* **57**, 2245–2262 (1973).
13. Montadert, L. et al. *Maurice Ewing Symp. Ser.* **3**, 154 (1979).
14. Savelli, C. et al. *J. Volcan. Geotherm. Res.* **5**, 257 (1979).
15. Vardabasso, S. *56^e Congr. Soc. Geol. It., Guida alle escursioni*, 1–75 (1952); *Soc. Geol. It.* **21**, 159–179 (1955).
16. Chabrier, G. *C. R. heb. Séanc. Acad. Sci., Paris* **271**, 1251–1255 (1970).
17. Letouzey, J., Wannesson, J. & Chérchi, A. *C. R. heb. Séanc. Acad. Sci., Paris* **294**, 596–602 (1982).
18. Barca, S. & Palmerini, V. *Boll. Soc. Sarda Sc. Nat.* **12**, 13–50 (1973).
19. Chérchi, A. & Schroeder, R. *Bull. Soc. géol. Fr.* **18**, 1217–1219 (1976).
20. Chérchi, A. *Riv. Ital. Paleont. Strat.* **35**, 353–410 (1979).
21. Chérchi, A. & Schroeder, R. *Boll. Soc. Sarda Sc. Nat.* **20**, 27–36 (1981).
22. Pittau, P. *Boll. Soc. Paleont. It.* **18**, 303–314 (1979).
23. Pecorini, G. & Chérchi, A. *Mem. Soc. geol. It.* **8**, 421–451 (1969).
24. Mattauer, M. *C. R. Som. Soc. géol. Fr. No.* **10**, 316–317 (1962).
25. Chérchi, A. *V Congr. Néogène Méditerran. Lyon 1971, BRGM 78*, 433–445 (1974).
26. Wright, R. *Init. Rep. DSDP Leg 42A*, 837–844 (1978).
27. Arthaud, F., Ogier, M. & Seguret, M. *Bull. BRGM 1.3.*, 175–193 (1980–1981).
28. Bellaiche, G. et al. *Mar. Geol.* **16**, 47–56 (1974).
29. Bellaiche, G. et al. (Groupe Cyaligire) *Bull. Soc. géol. Fr.* **21**, 533–543 (1979).
30. Bourrouilh, R. thesis, Univ. Paris VI (1974).
31. Colom, G. *Rev. R. Acad. Cienc. Ex. Fis. nat. Madrid* **70**, 353–408 (1976).
32. Pomar, L. & Colom, G. *Bol. Soc. Hist. Nat. Baleares* **22**, 119–136 (1977).
33. Deriu, M. *Mem. Soc. Geol. It.* **3**, 675–706 (1962).
34. Maccioni, L. in *Paleogeografia del Terziario sardo nell'ambito dell Mediterraneo occidentale* (ed. Chérchi, A.) 277–282 (1974).
35. Coulon, C. thesis, Univ. Marseille III (1977).
36. Montigny, R., Edel, J. B. & Thuizat, R. *Earth planet. Sci. Lett.* **54**, 261–271 (1981).
37. Rehaut, J. P. thesis, Univ. Paris VI (1981).
38. Vail, P. R. & Hardenbohl, J. *Oceanus* **22**, 71–79 (1979).

High rock stresses as a consequence of glaciation

Anders Carlsson

Swedish State Power Board, S-162 87 Vallingby, Sweden

Tommy Olsson

K-Konsult, S-117 80 Stockholm, Sweden

Knowledge of the stress field in the Earth's crust is crucial for the understanding of processes such as faulting and fracturing. It is also of fundamental importance for the design and construction of underground structures. The two most common methods used for *in situ* stress determinations for engineering geology purposes are (1) the stress relief method, in which the differences in strain resulting from induced stress relaxation are studied, and (2) the hydraulic fracturing method, in which one of the presumed principal stresses is replaced by a hydraulic pressure. Of these, the stress relief method is more applicable to design and construction problems as it enables the three-dimensional stress tensor to be measured. The *in situ* rock stress measurements reported here were carried out in a 500-m deep borehole in Precambrian bedrocks in Sweden and show high stress levels throughout the rock mass. A comparison between the actual recorded stresses and the theoretical stress field from a 2,000 m-thick land ice may provide a possible explanation for the measured stress situation.

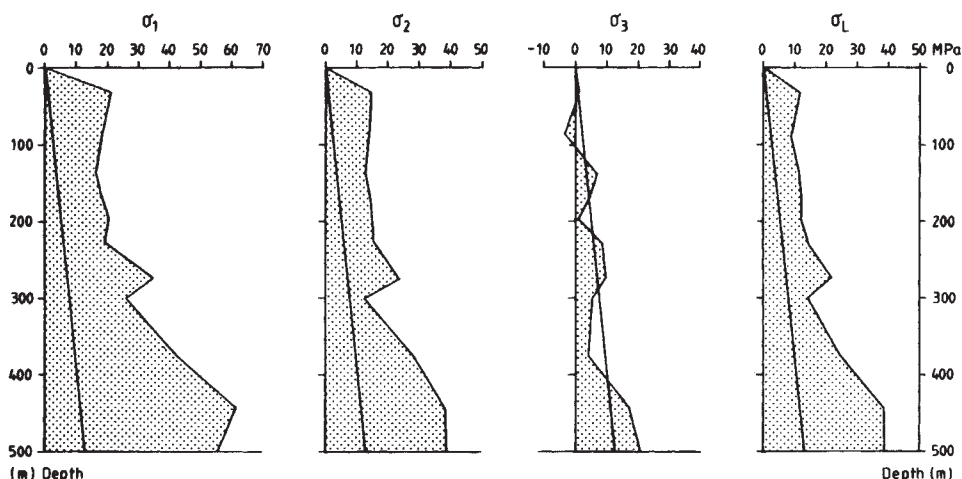


Fig. 1 Principal stresses (σ_1 , σ_2 and σ_3) versus depth obtained in borehole DBT1. σ_1 and σ_2 are both nearly horizontal and σ_3 is nearly vertical. The stress level (σ_L), defined as the mean value of the principal stresses, is also shown. The straight lines represent the lithostatic stresses given by the weight of the overburden.

The only one stress relief device that we know of that is suitable for measurements in deep, water-filled boreholes has been developed at the Swedish State Power Board¹. It is based on Leeman's three-dimensional theories and has been used in Sweden during the past 5 years where recordings have been taken at different locations in virgin parts of the Swedish bedrock.

Measurements have been made in several boreholes drilled from the ground surface, with readings taken at certain intervals down the hole. In the Forsmark area in southern, central Sweden, 30 successful measurements, concentrated to 11 levels, were taken during drilling of a 503 m-deep borehole (borehole DBT1). Another five boreholes have been studied in this area, but with fewer measurements and at shallower depths.

The magnitudes of the principal stresses at each of the 11 levels are shown in Fig. 1 as well as the stress level, defined as the mean value of the principal stresses. Figure 2 indicates the distribution of the orientation of the principal stresses. As may be seen, the stresses are generally high, compared with the weight of the overburden given by the straight line in Fig. 1. The maximum principal stress is ~70 MPa and acts in a plane near the horizontal. The dominating direction of the maximum

principal stress (σ_1) is roughly WNW-ESE in the uppermost 300 m, below which level it shifts to a more NW-SE direction. This is indicated by the clockwise arrow in Fig. 2. The dip is approximately horizontal. The smallest of the principal stresses (σ_3) is approximately vertical, acting in a direction that dips steeply to the south. Note that there is a marked increase in the magnitude of stresses below 300 m, which coincides with the observed shift in direction.

The geology of Sweden is characterized by a basement of crystalline rocks of Svecocarelian age (1,500 Myr), which is covered by a thin layer of Quaternary deposits, mainly till. Sweden is located on the western edge of the Eurasian Precambrian shield, bounded to the west by the Caledonian Mountain Range, which forms the border between Norway and Sweden. In large areas, such as Forsmark, the crystalline basement has been eroded down to a peneplain of Precambrian age.

The bedrock at Forsmark is composed chiefly of gneiss granites, mica gneiss and mica schists. The rock stress measurements were performed in the gneiss granitic part of the area.

As shown in Fig. 1, the principal stresses found are of a higher magnitude than would be expected if the stress field were governed by the weight of the overburden alone. This is

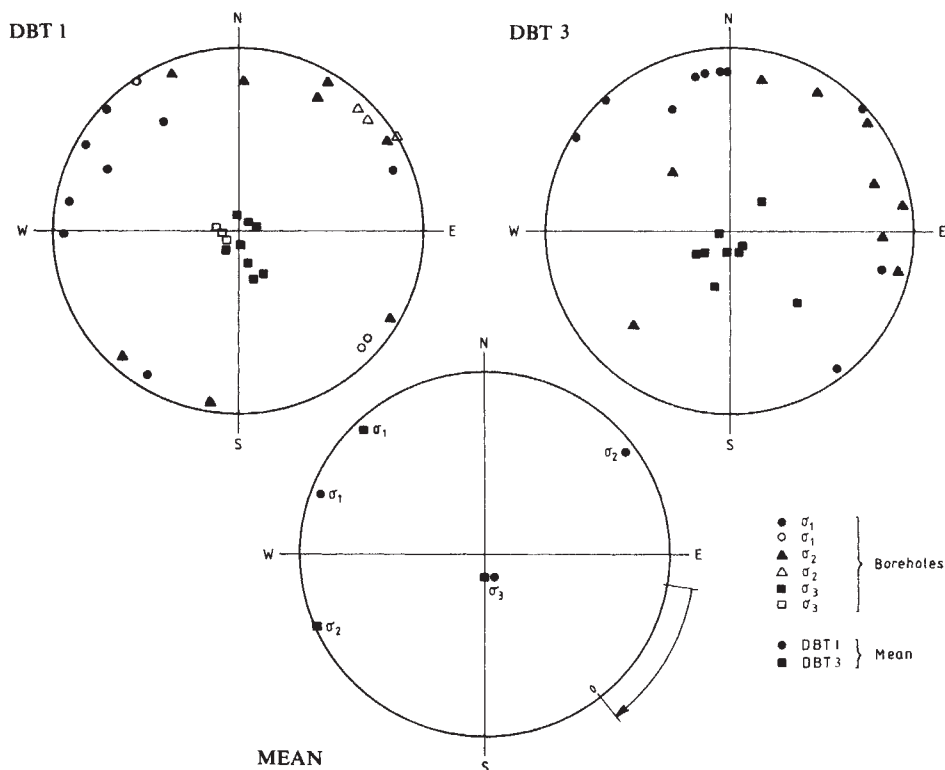
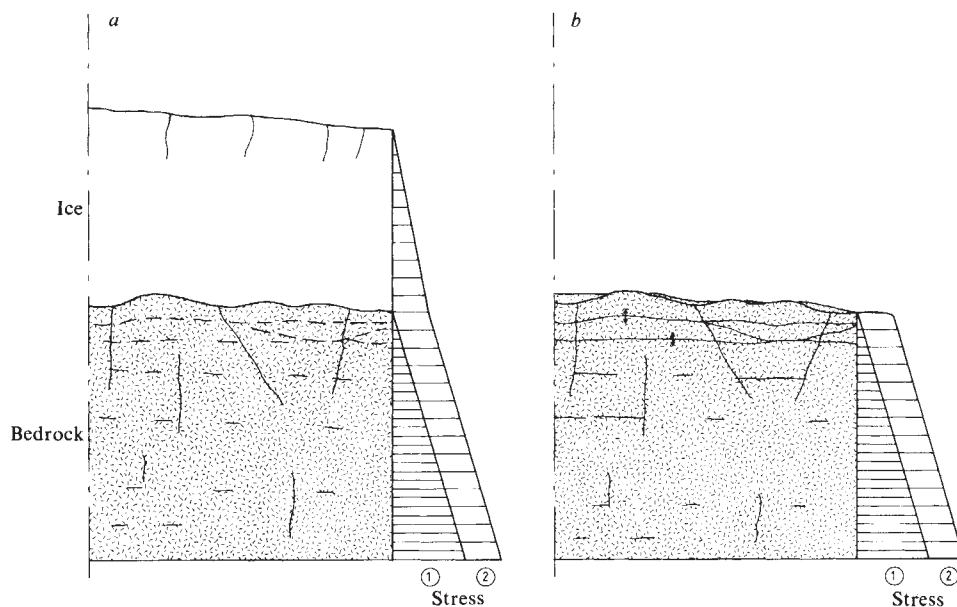


Fig. 2 Orientation of the individual principal stresses and mean values for the boreholes DBT1 and DBT3. The orientations are given as equal-area plots of the intersection of the principal axes with the lower hemisphere. The clockwise shift indicated by the arrow in the lower figure denotes a shift in direction at the passage of the fracture zone at 320 m. Filled and open symbols in DBT1 represent recordings taken above and below the 320-m depth, respectively.

Fig. 3 Hypothetical influence of glaciation. The stress level: *a*, during glaciation; *b*, after deglaciation. Stresses 1 and 2 denote the gravitational stresses due to bedrock and ice respectively.



the case for all principal stresses including the nearly vertical one (σ_3). For the vertical component, the correspondence between the measured stress and the lithostatic is fairly good in the uppermost 300 m, but below this depth, the vertical component is much higher than the theoretically expected value. It is therefore not possible to explain the high stresses which actually exist as gravitational stresses; they must be of some other origin.

A high stress level in the bedrock may be due to many factors. An explanation frequently proposed is that denudation of the rock surface has occurred and this has decreased the weight of the overburden significantly. The lateral stress field created by the earlier formation should then remain, as long as the strength of the rock is not exceeded (allowing for an elastic adjustment of lateral stresses due to vertical stress decrease).

Voight² and others have chosen this theory to explain anomalous stress situations. This theory may be applicable to many areas, but not to Forsmark. Here the major denudation occurred in the Precambrian and it is not reasonable to believe that the present stress situation should have its origin back in this age.

Jamison and Cook³ have analysed many rock stress measurements, and concluded that stresses may be influenced by the

frictional resistance to sliding between individual blocks. This explanation is suitable for the stress situation obtained, but it does not provide any information on the origin of the stresses in the Forsmark area, sited as it is on the Precambrian peneplain, on a stable shield of old, crystalline rocks. Other explanations must be sought for the measured magnitude of the stress field at Forsmark. Comparing the results observed with theoretical stresses that could be expected to prevail during a glaciation, it is possible to give a theory for the state of stress in the Forsmark area. Deglaciation is a very rapid process (in the geological time scale) which may be compared with a very rapid denudation, and thus accords with the Voight hypothesis.

At the glaciation maximum the surface of the underlying bedrock is exposed to a vertical stress determined by the weight of the ice cap. By adopting Voight's theories, it is seen that the ice also creates lateral horizontal stresses initially of the magnitude of one-quarter of the vertical stress. This is based on a Poisson's ratio of 0.2 for the rock mass at Forsmark (33 determinations in the range 0.17–0.23). The lateral stress will increase in time due to creep to a value higher than a quarter of, but lower than, the vertical stress. During deglaciation, the vertical stress falls elastically down to zero, while the horizontal stress only falls by one-quarter. Thus it is possible to assume

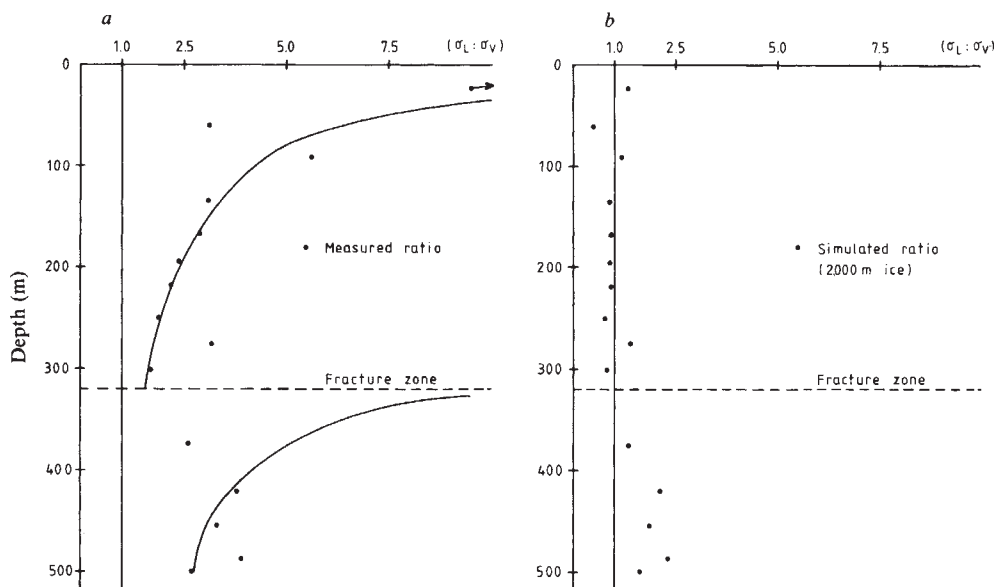


Fig. 4 The ratio between measured (σ_L) and theoretical (σ_V) stress level plotted against depth. *a*, The ratio between the recorded values and the theoretical as it is today. *b*, The ratio between the recorded values and a theoretical value given by the weight of a 2,000 m-thick ice and the weight of the rock mass.

that the glaciation could have created a stress situation as shown in Fig. 3.

To test this theory, the present results of stress measurements have been compared with theoretical effects of an assumed ice cap, 2,000 m thick. The result of the comparison is shown in Fig. 4, where Fig. 4a plots the ratio between the actual measured stress level (mean stress) and the theoretical lithostatic stress field ($\sigma_L:\sigma_v$ ratio) against depth. This curve indicates that the discrepancy between the measured stress level and the theoretical level decreases with increasing depth. Simulating the pressure from a 2,000 m-thick ice cap gives a stress level of 20 MPa at the bedrock surface, in lithostatic conditions.

Using the sum of this ice-induced stress and a lithostatic stress level in the rock due to the weight of the overlying rock mass to give a new value for σ_v , it is possible to recalculate the $\sigma_L:\sigma_v$ ratio, which includes the effects of a glaciation. Figure 4b shows that the theoretical, lithostatic stress level equals the measured one. Hence the present stress level, as measured at Forsmark, may be an effect of Quaternary glaciations, where the deglaciation resembles a rapid denudation.

However, the glaciation cannot be the sole cause of the high stresses recorded below a fracture zone at 320 m indicated in Fig. 4. Below this level readings of about 70 MPa for the maximum principal stress were taken. These high stresses could be explained as being stresses of palaeotectonic origin.

The effect of a glaciation on the stress field is discussed more completely in ref. 4.

Received 22 December 1981; accepted 9 June 1982.

1. Hiltcher, R., Martna, J. & Strindell, L. *Proc. 4th int. Congr. Rock Mech.* **2**, 227–234 (1979).
2. Voight, B. *Proc. 1st int. Congr. Rock Mech.* **2**, 45–50 (1966).
3. Jamison, O. B. & Cook, N. G. W. *J. geophys. Res.* **85**, B4, 1833–1838 (1980).
4. Carlsson, A. & Olsson, T. *Swedish State Power Board, Research and Development Rep.* 5:1 (1982).

Evidence for $^{13}\text{C}/^{12}\text{C}$ fractionation between tree leaves and wood

Steven W. Leavitt & Austin Long

Laboratory of Isotope Geochemistry, Department of Geosciences, University of Arizona, Tucson, Arizona 85721, USA

An isotopic fractionation step between carbon fixed in the leaves and that assimilated into wood cellulose has not previously been included in atmosphere–plant fractionation models. Data presented here, however, show that the magnitude of such a fractionation step may be significant. In an experiment which involved 10 juniper trees from around Arizona, cellulose of leaves was found to be isotopically lighter than that of the corresponding tree rings in all trees by an average of 2‰. In a second experiment, the $\delta^{13}\text{C}$ of leaf material from a juniper tree sampled at monthly intervals was compared with the $\delta^{13}\text{C}$ of the corresponding ring. For cellulose, $\delta^{13}\text{C}$ of the wood varied with changes in $\delta^{13}\text{C}$ of the leaves and was persistently isotopically heavier by 3–4‰. Furthermore, the observed change in $\delta^{13}\text{C}$ through the growing season suggests a temperature coefficient of about $-0.27\text{‰}^\circ\text{C}^{-1}$. The direction of this leaf–wood fractionation precludes diffusion processes as the primary cause.

For the most part, the carbon assimilated into woody tissue of trees is initially fixed from atmospheric carbon dioxide by chlorenchyma of the leaves. This process serves as the basis for using tree rings to decipher past changes in atmospheric radiocarbon production or dilution¹, and more recently to determine the role of the biosphere as a CO_2 source or sink from tree-ring $\delta^{13}\text{C}$ measurements². The $\delta^{13}\text{C}$ fractionation between atmospheric CO_2 and C_3 plants, including most trees, is typically about -18‰ , the greatest portion of which takes place at the

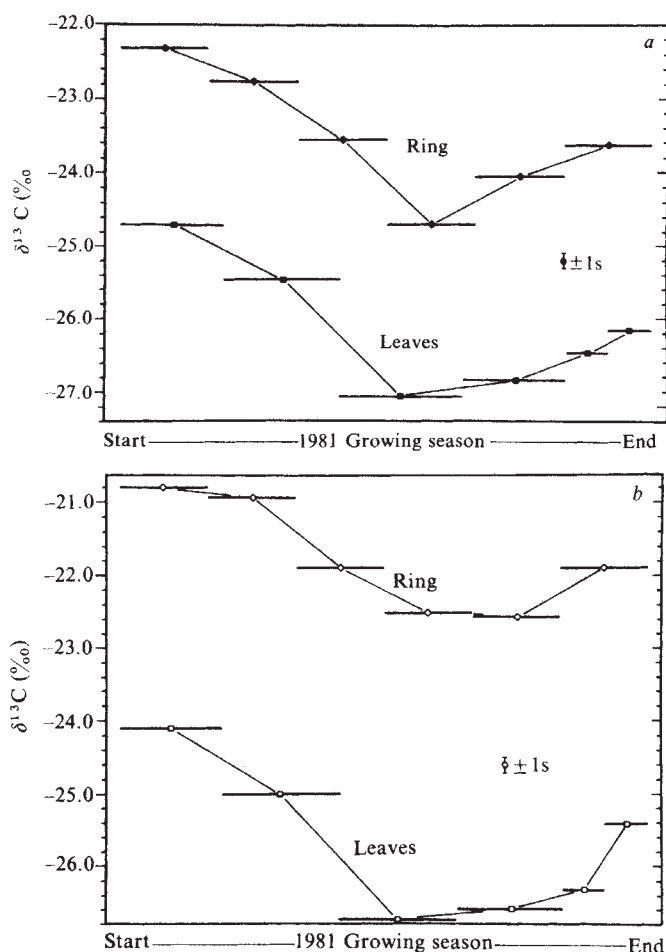


Fig. 1 a, Whole-tissue $\delta^{13}\text{C}$ analyses; b, cellulose $\delta^{13}\text{C}$ analyses of *J. monosperma* ring subdivisions and leaves from the 1981 growing season. Positions of leaf analyses in time have been scaled to their proportion of the total seasonal growth (Table 1).

enzyme fixation step. However, models which attempt to account for the atmosphere–plant fractionation either do not consider isotopic fractionation after glucose formation^{3,4}, or consider any fractionation during the formation of other compounds (including cellulose in tree rings) from photosynthesized sugars to be relatively slight or negligible^{5,6,20}.

Evidence, however, suggests isotopic selection during formation of individual organic compounds within plants. For example, a range of several per mil $\delta^{13}\text{C}$ among different organic constituents is well documented^{7,8}. Analysis of whole tissue indicates that leaves are lighter than wood by 1–2‰^{6,9,11}, which implies isotopic differences between tissues also. Unfortunately, this evidence of leaf–wood $\delta^{13}\text{C}$ fractionation is inconclusive because of the possibility that such whole-tissue differences may simply reflect varied proportions of organic compounds, each of which has a different isotopic composition. Therefore, analyses of individual compounds such as cellulose are necessary to test more conclusively for this fractionation. So far, such information is scarce. One study (A. L. and J. C. Lerman, unpublished data) examined $\delta^{13}\text{C}$ of cellulose from leaves and wood of several bristlecone pine individuals and found in most cases that the leaf cellulose was isotopically lighter. Unfortunately, although the leaves were compared with wood of the same sites and of the same age, the leaves and wood were not from the same individual trees.

We conducted two experiments which provide insight into the nature of fractionation between leaves and wood. At the end of the 1979 growing season we sampled juniper trees in Arizona growing at elevations ranging from 1,200 to 2,300 m. Leaf samples were pooled from the four cardinal directions and

Table 1 $\delta^{13}\text{C}$ values of cellulose from leaves and wood of Arizona juniper trees harvested after 1979 growing season

Site	Species	Elevation (m)	δ^{wood} (1970–79) (‰)	δ^{leaves} (‰)	$\Delta\delta^{\text{wood-leaves}}$ (‰)
Globe	<i>Juniperus monosperma</i>	1,195	–19.3	–21.6	+2.3
Jerome	<i>J. monosperma</i>	1,845	–21.6	–22.8	+1.2
Oracle	<i>J. deppeana</i>	1,450	–18.9	–22.4	+3.5
Prescott	<i>J. deppeana</i>	1,705	–20.8	–22.2	+1.4
St Johns	<i>J. monosperma</i>	1,790	–20.4	–21.7	+1.3
Santa Rita	<i>J. deppeana</i>	1,370	–20.8	–22.3	+1.5
Seligman	<i>J. osteosperma</i>	1,605	–21.0	–23.0	+2.0
Snowflake	<i>J. monosperma</i>	1,735	–20.0	–22.8	+2.8
Springerville	<i>J. monosperma</i>	2,315	–20.9	–22.9	+2.0
Williams	<i>J. osteosperma</i>	2,055	–20.3	–22.8	+2.5

$\bar{x} = +2.0$; $s = 0.7$.

a section of wood was taken from the base of the tree. Because junipers retain ~10 yr of leaves at any time¹², the rings representing 1970–79, taken about their full circumference, were collected for comparison. Cellulose from the leaf and wood samples was isolated using an acidified sodium chlorite solution in a procedure modified after Green¹³. After combustion of the samples to CO_2 , carbon isotopic analyses were performed with a Micromass 602C mass spectrometer. The results reported in Table 1 are expressed in the standard δ notation with respect to the PDB standard¹⁴, with the reproducibility of combustion and analysis estimated to be $\pm 0.1\%$. The $\delta^{13}\text{C}$ of cellulose in leaves in all cases was lower than the corresponding wood by an average of $2.0 \pm 0.7\%$.

In the second experiment, we harvested leaf material at approximately monthly intervals over the 1981 growing season from a juniper tree (*Juniperus monosperma*) in the grounds of the University of Arizona. Marking of leaf tips along one major branch at the beginning of the intervals ensured that only the growth for that particular interval would be harvested. Furthermore, marking a large number of tips allowed some to be harvested while others could be remarked for harvesting later. Additional leaf tips, not previously marked, were also marked at the beginning of new intervals. These precautions minimized the effects that the cutting may have had on growth and provided a large pool of leaf tips for cutting at the end of each interval.

Table 2 summarizes the harvesting schedule and average amount of growth for only those tips which were harvested. The average growth values for the 6 August–7 September and the 7 September–4 October periods are higher than the actual growth of all tips during those two intervals because ~20% and 50%, respectively, of the tips marked at the beginning of each of these intervals did not grow at all and were not included in the averages. Finally, on 4 October, several tree-ring cores were taken from the trunk and branch. The 1981 ring from one core was divided into six approximately equal sections and cellulose was isolated from splits of these ring subdivisions and from the leaf samples.

Results of carbon isotopic analysis of the whole-tissue and the cellulose component of these samples are shown in Fig. 1. Strictly speaking, direct comparison of the wood and leaves is difficult because the time period represented by each of the wood subdivisions is unknown. However, as a first approximation we have plotted the wood analysis as equal lengths in time and then scaled the position of each leaf analysis to its proportion of the overall seasonal leaf growth. The major result is that leaves are persistently lighter than wood for both whole-tissue and cellulose by ~2.5 and 3.5‰, respectively, although leaf whole-tissue and cellulose values are quite similar. A pattern of decreasing $\delta^{13}\text{C}$ is evident in the spring and early summer followed by increasing values for the remainder of the sampling periods. The magnitude of the $\delta^{13}\text{C}$ decrease cannot be explained by the seasonal variation of $\delta^{13}\text{C}$ of atmospheric CO_2 which only changes ~0.4‰ in this region¹⁵. Moisture may be considered non-limiting because this tree is well irrigated, however, allowing us to calculate a significant temperature response of -0.27% $\delta^{13}\text{C}$ $^\circ\text{C}^{-1}$ ($r^2 = 0.86$). This coefficient results from regression of the April to September mean monthly

Table 2 Leaf sampling schedule for the *J. monosperma* individual

Harvest date	No. of leaves sampled	Mean net growth (mm)	s.d. (mm)
6 May 1981	28	5.1	1.8
7 June 1981	34	5.6	2.2
9 July 1981	43	6.1	3.1
6 August 1981	37	5.1	2.0
7 September 1981	43	2.1	1.1
4 October 1981	35	2.0	1.2

temperatures for 1981 (9.7, 13.6, 18.9, 22.7, 20.2 and 20.4°C , respectively) with their associated leaf cellulose $\delta^{13}\text{C}$ results (–24.1, –25.0, –26.7, –26.6, –26.2 and –25.4‰, respectively). This coefficient is opposite in sign to that calculated in an intra-annual study of *Pinus radiata* tree rings from New Zealand¹⁶. However, it is similar in sign and magnitude to temperature coefficients from several *in vitro* studies with the carbon-fixing enzyme RuBP-C^{17,18}.

If we infer similar leaf–wood isotopic differences for other species and locations, these results indicate that detailed models for atmosphere-to-plant carbon isotopic fractionation will need to include an additional fractionation step of +2‰ or more between the leaves and wood. This fractionation probably cannot be explained by diffusive processes because of the large molecular weight of glucose and because diffusion away from the leaves would tend to enrich them in ^{13}C , just the opposite of these results. If some type of ^{12}C preference operates during assimilation, similar to the preference for ^{12}C in carbon fixation, then this may be manifested in a $\delta^{13}\text{C}$ change in cellulose away from the photosynthate source. We had hoped to test this explanation of an isotopic gradient from leaves to wood with the other cores, but unfortunately they did not display a clear ring pattern with which to distinguish the 1981 growth-ring boundaries. One previous study¹⁹ of longitudinal variation of $\delta^{13}\text{C}$ in *Quercus rubra* found a fairly constant $\delta^{13}\text{C}$ in a ring when measured over the relatively short distance of 40 cm.

Received 5 March; accepted 17 May 1982.

- Damon, P. E., Lerman, J. C. & Long, A. A. *Rev. Earth planet. Sci.* **6**, 457–494 (1978).
- Stuiver, M. *Science* **199**, 253–258 (1978).
- Troughton, J. H. in *Proc. 8th int. Conf. Radiocarbon Dating*, E39–E57 (Royal Society of New Zealand, Wellington, 1972).
- Grinstead, M. J. thesis, Univ. Waikato (1977).
- Park, R. & Epstein, S. *Geochim. cosmochim. Acta* **21**, 110–126 (1960).
- Francey, R. J. in *Carbon Dioxide and Climate: Australian Research* (ed. Pearman, G. I.) 95–104 (Australian Academy of Science, Canberra, 1980).
- Park, R. & Epstein, S. *Pl. Physiol.* **36**, 133–138 (1961).
- Degens, E. T., Guillard, R. R. L., Sackett, W. M. & Helleburst, J. A. *Deep-Sea Res.* **15**, 1–9 (1968).
- Craig, H. *Geochim. cosmochim. Acta* **3**, 53–92 (1953).
- Lerman, J. C. in *Proc. 8th int. Conf. Radiocarbon Dating*, H16–H28 (Royal Society of New Zealand, Wellington, 1972).
- Stout, J. D. & Rafter, T. A. *DSIR Bull.* **220**, 75–83 (1978).
- Arnold, L. D. thesis, Univ. Arizona (1979).
- Green, J. W. in *Methods of Carbohydrate Chemistry III* (ed. Whistler, R. L.) 9–21 (Academic, New York, 1963).

14. Craig, H. *Geochim. cosmochim. Acta* **12**, 133–149 (1957).
15. Keeling, C. D. *Geochim. cosmochim. Acta* **24**, 277–298 (1961).
16. Wilson, A. T. & Grinstead, M. J. *Nature* **265**, 133–135 (1977).
17. Christeller, J. T., Laing, W. A. & Troughton, J. H. *Pl. Physiol.* **52**, 580–582 (1976).
18. Estep, M. F., Tabita, F. R., Parker, P. L. & Van Baalen, C. *Pl. Physiol.* **61**, 680–687 (1978).
19. Tans, P. P. & Mook, W. G. *Tellus* **32**, 268–283 (1980).
20. Francey, R. J. & Farquhar, G. D. *Nature* **297**, 28–31 (1982).

Oxic diagenetic products as indicators of past oceanic algal populations and diversity

F. T. Gillan* & R. B. Johns

Department of Organic Chemistry, University of Melbourne, Parkville, Victoria 3052, Australia

Marine sediments contain a great diversity of organic molecules which have originated *in situ* and from planktonic organisms. The establishment of a direct relationship between possible 'chemical markers' and source organisms would enable studies of biomass-turnover time to be undertaken, using dated sediment cores as temporal records. Studies of the influence of long-term environmental variations on the abundance of planktonic flora and fauna are potentially amenable to this approach. Chemical compounds biosynthesized by organisms in the water column are, however, subjected to microbial attack, oxidation and other degradative processes before, and indeed after, sedimentary incorporation¹. We show here that changes in the sedimentary abundances and distribution of selected chemical moieties produced by oxidative degradation of fatty acids of planktonic origin, can provide evidence for a change in algal abundance and distribution in the overlying oceanic waters.

As part of a continuing study² of the geochemistry of Patton escarpment sediments (32°34'.4 N, 120°28'.0 W; depth: 3,800 m), a sediment core has been analysed in detail. Four samples (1), 0–5 cm; (2), 12.5–17.5 cm; (3), 27.5–32.5 cm; (4), 42.5–47.5 cm) from the core were solvent-extracted and the residues from the extraction were saponified. Selected lipid classes from the bound acidic fraction were examined (after derivatization as the methyl esters, trimethylsilyl ethers) by capillary GC on a SCOT SE30 column, and by GC/MS (HP 5995A GC/MS with fused silica SE30 WCOT capillary column). α,ω -Dicarboxylic acids ('diacids') were readily identifiable by their characteristic mass spectra³. A group of compounds with a strong m/z 175 in their mass spectra, which did not co-elute with β -hydroxy acids was also observed. These compounds were not observed unless the methyl esters were trimethylsilylated. The mass spectrum of the major component was m/z (I): 73(100); 75(62); 89(77); 133(54); 159(31); 175(45); 261(96) and 292(58). The ions: m/z 73, 75, 89, 133, 159 and 175 are indicative of a β -trimethylsilyl ether-methyl ester, while the intense M-31 fragmentation suggests an α,ω -dimethoxyl function in the molecule. The m/z 75: m/z 73 ratio is higher than we have observed with typical β -hydroxy acid derivatives and probably indicates some contribution to m/z 75 from the characteristic m/z 75 fragmentation of a dimethyl acetal⁴. The most probable structure for this compound is the dimethyl acetal, trimethylsilyl ether, methyl ester derivative of β -hydroxy- ω -oxo-heptanoic acid: such compounds have not previously been reported in sediments. No authentic sample of a compound of this type was available for mass spectral comparison.

Table 1 lists the abundances of bound diacids in the four sediment samples examined. Note that the bound diacid abundance is much greater than that of any other lipid fraction we have determined (for example, solvent-extractable monocar-

Table 1 Depth profiles of bound dicarboxylic acids in the Patton escarpment sediments

Diacid	Abundance (p.p.m. wet wt) [†]			
	(1)	(2)	(3)	(4)
7:0*	0.73	1.86	3.3	2.4
8:0	1.40	1.28	2.5	1.60
9:0	10.9	11.4	13.5	10.9
10:0	0.45	0.42	0.75	0.65
11:0	0.42	0.55	1.19	0.94
12:0	0.043	0.058	0.128	0.099
13:0	0.015	0.041	0.066	0.032
14:0	0.014	0.027	0.052	0.017
15:0	0.033	0.035	0.059	0.020
16:0	0.025	0.031	0.049	0.017
16:1	0.016	0.010	0.007	0.001
17:0	0.014	0.023	0.037	0.018
18:0	0.016	0.023	0.050	0.027
18:1	0.017	0.002	0.001	0.001
19:0	0.023	0.022	0.052	0.029
20:0	0.025	0.021	0.064	0.031
21:0	0.019	0.017	0.050	0.026
22:0	0.013	0.016	0.035	0.022
23:0	0.017	0.016	0.034	0.024
24:0	0.014	0.018	0.038	0.027
25:0	0.012	0.013	0.029	0.025
26:0	0.005	0.013	0.024	0.022
27:0	0.001	0.011	0.014	0.016
Σ	14.2	15.9	22.1	16.9

* Diacids are denoted by $M:N$, where M is the number of carbon atoms and N the number of double bonds.

[†] All sediments 68 ± 2% moisture.

boxylic acids: 6.5 p.p.m., hydrocarbons (aliphatic): 1.3 p.p.m.)². The main component in each sample is the C₉ diacid (9:0), which constitutes ~70% of the total observed diacids; 7:0, 8:0, 10:0 and 11:0 diacids are also major components. The longer chain-length (>C₁₁) diacids constitute only a small percentage of the total diacid abundance (~3%). In contrast, in an anaerobic deep-sea sediment from the Santa Catalina Basin, bound 9:0 diacid is a minor constituent (<2% of the diacids) (our unpublished results). This indicates that the short-chain diacids probably originate by oxidation of lipids input to the oxic sediment. The high abundance of the C₉ diacid precludes a major contribution from, for example, higher plant sources. The production of similar diacid distributions to those observed here, in oxidations of algal kerogens⁵, supports the probable origin of the short-chain diacids by oxidative degradation of suitable organic substrates present in the lipid mixture deposited in these sediments. Δ^9 monoenoic, monocarboxylic acids are major components of the lipids of marine sediments, 16:1 Δ^9 being the most abundant monoenoic acid in many sediments we have studied⁶. Oxidation of this acid could lead to a variety of products (see Table 2). Direct cleavage at the double bond would produce only the C₉ diacid, while hydration followed by oxidative cleavage either side of the hydroxyl would lead to a mixture of C₈, C₉ and C₁₀ diacids. ω -Oxidation (to Di-16:1 Δ^7)

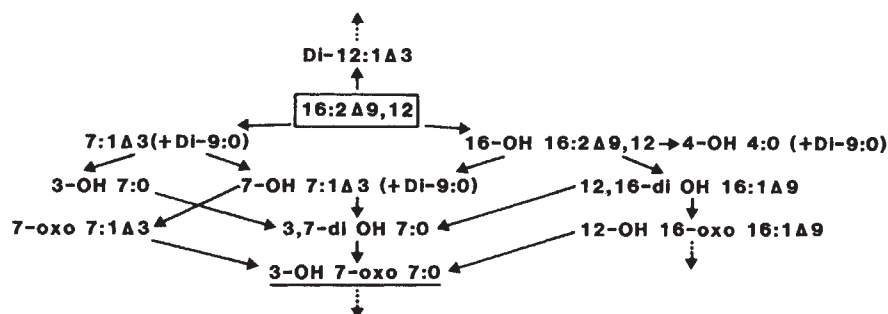
Table 2 Expected diacid products from oxic diagenesis of palmitoleic acid (16:1 Δ^9)

Process	Diacid products*				
	6:0	7:0	8:0	9:0	10:0
Direct cleavage	0	0	0	1	0
Hydration then cleavage*	0	0	1	2	1
ω -Oxidation then cleavage	0	1	0	1	0
ω -Oxidation, hydration then cleavage	1	1	2	1	1

* These product ratios assume that hydration will produce equal amounts of the 9-OH and 10-OH products and cleavage of a hydroxy-acid will occur equally on either side of the hydroxyl-carbon atom (for example, for 9-OH 16:0 FA, C₉/C₁₀ and C₈/C₉ cleavages are equally probable).

* Present address: Australian Institute of Marine Science, Townsville MSO, 4810, Queensland, Australia.

Fig. 1 Oxic diagenesis of a common, diatom-derived dienoic acid.



before oxidative cleavage is also possible. The observed diacid distribution indicates that direct cleavage is the predominant process, and the much higher abundance of C_8 diacid relative to the C_{10} diacid supports some contribution from the ω -oxidation, hydroxylation and then cleavage pathway (the alternate pathway: hydration then cleavage would result in a ~1:1 C_8 : C_{10} ratio). The other major diacid, Di-11:0, will probably result from oxidation of $\Delta 11$ monoenoic acids. These acids are the second most abundant group of monoenoic acids in marine sediments⁶. Oxidation of the $\Delta 11$ acids would result in the production of a small quantity of Di-12:0. This proposed interrelationship of C_{11} and C_{12} diacids is supported by the high correlation of these data ($R^2 = 0.9996$). Whilst ω -oxidation is a known microbial process, microbial degradation of fatty acids generally proceeds through a series of both β -oxidations and ω -oxidations⁷. Reactions such as this would not lead to the observed diacid distribution. No evidence for a microbially-mediated direct cleavage or hydration/cleavage process has been reported: the major degradative pathway is probably abiological. The observed diacid distribution is thus consistent with an input to the sediment of $\Delta 9$ (probably 16:1 $\Delta 9$ and 18:1 $\Delta 9$) and $\Delta 11$ (probably 18:1 $\Delta 11$) monocarboxylic acids, and subsequent oxidation.

As indicated above, the monoenoic acid distributions of intertidal marine sediments are dominated by 16:1 $\Delta 9$, 18:1 $\Delta 9$ and 18:1 $\Delta 11$ (ref. 6). These acids are believed to originate from *in situ* algae. In the Patton escarpment sediment there is no chemical evidence for the presence of *in situ* algae (algal carotenoids are not detectable²), however, particulates derived from planktonic algae are expected to contribute to the lipids of the sediments. The presence of typical algal sterols in these sediments², in particular, supports the proposed contribution to the sedimentary organic matter from algal debris.

The long-chain diacids detected in these sediments exhibit fairly smooth distribution envelopes with Carbon Preference Indices (CPI) near unity. If these acids were derived from a

higher plant input the CPI would be expected to be large (>3) (ref. 8). Rather, these acids are probably diagenetic products possibly produced by microbial di-terminal oxidation of alkanes or ω -oxidation of long-chain fatty acids, and subsequent α - and β -oxidations. The traces of monounsaturated diacids detected could originate by ω -oxidation of the corresponding monocarboxylic acid, or may indicate a contribution to the sedimentary lipids from higher plant suberins⁸.

The abundances of the β -hydroxy ω -oxo acids are listed in Table 3. By analogy with the mode of formation of the major diacids (Table 2), one can envisage a series of double bond cleavages, terminal oxidations and hydrations that would lead to such acids from a dienoic acid starting material (see Fig. 1). Note that these polyfunctional acids would be intermediates in the overall degradative pathway: it is expected that they would not be very abundant, as observed. The major acid of this class, C_7 , could originate from 16:2 $\Delta 9, 12$ monocarboxylic acid, whilst the second most abundant component, C_9 , would be derived from dienoic acids such as 18:2 $\Delta 9, 12$ or 16:2 $\Delta 7, 10$. The major monoenoic and dienoic acids present in planktonic algae are algal class dependent. A change in the abundances of the different algal classes contributing to a sediment would result in a concomitant change in the input of unsaturated fatty acids, and thus in the abundances in the sediments of their diagenetic products.

Diatoms biosynthesize large amounts of 16:1 $\Delta 9$ and 16:2 $\Delta 9, 12$ compared with most other algae⁹, and would thus be expected to be major contributors to the abundances of the C_9

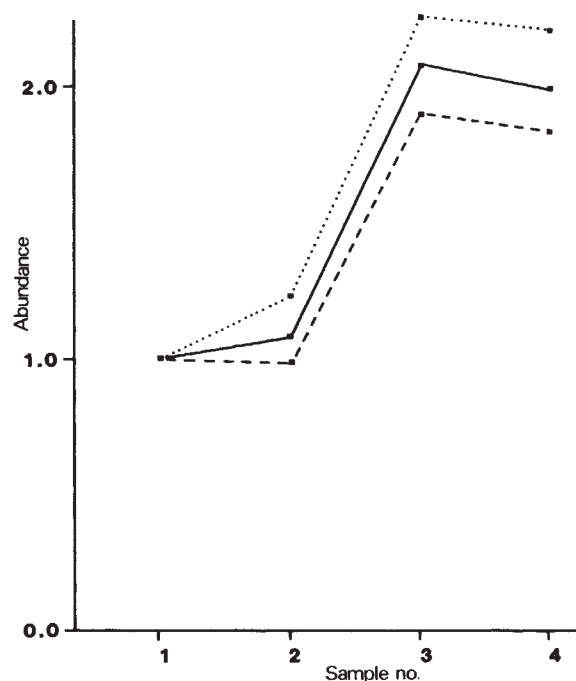


Fig. 2 Diacid ratios (C_{11}/C_9) and abundances of C_7 β -hydroxy- ω -oxo acid and 24-methylcholesta-5,22-dien-3 β -ol (all data normalized to 1.00 for sample (1)). ···, Diacid; ---, hydroxy acid; —, sterol.

Table 3 β -Hydroxy ω -oxo acids in the Patton escarpment sediments

Acid	Abundance (parts per 10^9 wet wt)			
	(1)	(2)	(3)	(4)
6:0	3	4	9	6
7:0	60	59	114	110
8:0	12	9	18	11
9:0	19	23	54	38
10:0	10	10	31	17
11:0	5	8	27	14
12:0	4	8	29	16
13:0	3	4	10	5
14:0	2	2	6	4
15:0	1	2	4	3
16:0	1	2	6	3
17:0	3	3	18	5
18:0	4	5	21	10
19:0	3	6	31	15
20:0	2	3	18	8
21:0	1	1	7	4
22:0	—	1	4	3
23:0	—	1	4	3
Σ	133	151	411	275

diacid and the C_7 β -hydroxy ω -oxo acid found in these sediments. In contrast, coccolithophores produce virtually no dienoic acids that would lead to the formation of the C_7 polyfunctional compound¹⁰; only C_9 precursors are present in these algae. The $\Delta 11:\Delta 9$ monoenoic acid ratios for the few algae of this class that have been analysed in detail also differ from those of diatoms. Similarly, certain green algae have been observed¹¹ to biosynthesize some 18:1 $\Delta 11$. Changes in the chain-length distributions of the diagenetically-derived acids could thus be indicative of changes in the algal populations contributing to the sedimentary lipids. 24-Methylcholesta-5,22-dien-3 β -ol is a common major sterol in coccolithophores¹⁰ although it also is found in diatoms¹³. A change in the diatom and/or coccolithophore abundance or distribution should result in a change in the abundance of this sterol.

The dicarboxylic acids (Table 1) exhibit a marked change in the $C_{11}:C_9$ ratio between samples (2) and (3) (ratios: (1):(2):(3):(4):0.039:0.048:0.088:0.086) due mainly to a large increase in the C_{11} diacid abundance. The abundance of β -hydroxy ω -oxo 7:0 similarly changes abruptly between

samples (2) and (3). Figure 2 shows a plot of these data together with the abundance of 'bound' 24-methylcholesta-5,22-dien-3 β -ol. The high correlation of the data (diacid:hydroxy-oxo-acid $R^2=0.973$; diacid:sterol $R^2=0.996$; hydroxy-oxo-acid:sterol $R^2=0.991$) supports the proposed algal origin for the acids and indicates a probable change in the algal abundance and/or distribution between deposition of samples (3) and (2). The sedimentation rate for these sediments has not been determined, though a rate in the range 0.5–2.0 cm kyr⁻¹ could be expected for this near-shore environment¹². It is tempting to speculate that it may be of the order of 1.5 cm kyr⁻¹: samples (3) and (4) would then have been deposited during the last ice age, at a time when the algal abundance and distribution would surely have differed from that at present.

We thank Dr A. Yayanos for the sediment core and Dr R. Esdaile and CSIRO Division of Materials Science for access to the HP 5995A GC/MS. F.T.G. acknowledges financial assistance from a CPG Research Award. We thank the secretarial staff at the Australian Institute of Marine Science for their assistance.

Received 8 April; accepted 29 June 1982.

1. Didyk, B. M., Simoneit, B. R. T., Brassell, S. C. & Eglinton, G. *Nature* **272**, 216–222 (1978).
2. Gillan, F. T., Johns, R. B. & Volkman, J. K. *Geochim. cosmochim. Acta* (submitted).
3. Ryhage, R. & Stenhagen, E. *Arkiv. Kemi* **14**, 483–496 (1959).
4. Beynon, J. K., Saunders, R. A. & Williams, A. E. *The Mass Spectra of Organic Molecules* (Elsevier, Amsterdam, 1968).
5. Philp, R. P. *Nature* **262**, 134–136 (1976).
6. Gillan, F. T. et al. in *Advances in Organic Geochemistry 1981* (in the press).

7. Doelle, H. W. *Bacterial Metabolism* 2nd edn (Academic, New York, 1975).
8. Kolattukudy, P. E., Croteau, R. & Buckner, J. S. in *Chemistry and Biochemistry of Natural Waxes* (ed. Kolattukudy, P. E.) 289–347 (Elsevier, Amsterdam, 1976).
9. Chuecas, L. & Riley, J. P. *J. mar. biol. Ass. U.K.* **49**, 97–116 (1969).
10. Volkman, J. K., Smith, D. J., Eglinton, G., Forsberg, T. E. V. & Corner, E. D. S. *J. mar. biol. Ass. U.K.* **61**, 509–527 (1981).
11. Johns, R. B., Nichols, P. D. & Perry, G. J. *Phytochemistry* **18**, 799–802 (1979).
12. Bostrom, K., Kraemer, T. & Gartner, S. *Chem. Geol.* **11**, 123–148 (1973).
13. Volkman, J. K., Gillan, F. T., Johns, R. B. & Eglinton, G. *Geochim. cosmochim. Acta* **45**, 1817–1828 (1981).

Age of the Lufeng, China, hominoid locality

Lawrence J. Flynn

Department of Geosciences, University of Arizona, Tucson, Arizona 85721, USA

Qi Guo-qin

Institute of Vertebrate Paleontology and Paleoanthropology, Academia Sinica, PO Box 643, Beijing (27), People's Republic of China

Recently discovered hominoids from a locality near Lufeng, China, are crucial to the understanding of the biogeographical distribution and the pattern of evolution of all later Miocene hominoids. However, the significance of these finds is difficult to assess fully without knowing their age accurately. The Lufeng fauna is diverse and includes several species of the rodent family Rhizomyidae. Rhizomyid evolution is well documented in the Siwalik deposits of Pakistan and provides the most precise method available at present for dating the Lufeng fauna. Based on co-occurrence of *Brachyrhizomys nagrii*, *Brachyrhizomys tetracharax* and cf. *Brachyrhizomys pilgrimi*, we conclude that hominoids inhabited southern China and Pakistan at about the same time, 8 Myr ago.

The success of expeditions in Pakistan¹ and Europe² in recovering late Miocene hominoids has stimulated fossil hunting in terrestrial deposits throughout southern Asia. Chinese scientists have recovered remarkably complete fossils from a site near Lufeng, Yunnan. Diverse rodents are frequently associated with hominoid remains and contribute to chronological correlation because evolution in some lineages was relatively rapid and is well documented. Fossil rhizomyids (bamboo rats of Asia and mole rats of Africa) are especially well represented in Pakistan³ and occur at Lufeng.

The Lufeng fossil locality is near Shihuiba, a small village ~9 km north of Lufeng city (102°4' E, 25°3' N). Within a Cenozoic sequence of fluvial-lacustrine deposits, fossils are concentrated in ~3 m of lignite in a small area⁴. This fauna is

diverse and includes abundant hominoid primates, numerous ungulates, and several carnivores and large rodents⁴. Unlike Siwalik faunas which span >10 Myr (Miocene–Pleistocene), the Lufeng fauna probably accumulated over a geologically short period of time. The Lufeng fauna resembles late Miocene faunal assemblages almost 3,300 km to the west in Pakistan, but differs from more northerly Chinese faunas⁵. Lufeng and the Pakistan sites apparently occupied the same general biogeographical province distinct from that of northern China in the late Miocene.

The Siwalik deposits of Pakistan are thousands of metres thick and vertebrates are widely distributed throughout the sequence. Palaeontological expeditions have recovered large collections of vertebrates, with documented stratigraphical provenance. Temporal calibration of Siwalik fossil horizons is based on palaeomagnetic reversal sequences and on several radiometric dates. Strata in the temporal interval from ~9 to 6.5 Myr are particularly well documented because the magnetic sequence includes many distinctive reversals, sedimentation was relatively continuous at this time, and magnetic properties of the sediment preserved the polarity sequence without ambiguity. Many palaeomagnetic sections through the 8-Myr

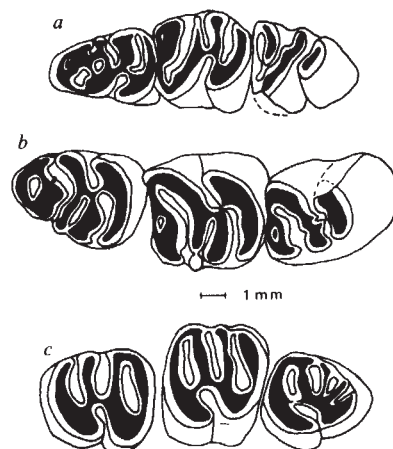


Fig. 1 Rhizomyid dentitions from Lufeng. a, *Brachyrhizomys nagrii*; b and c, *B. tetracharax*.

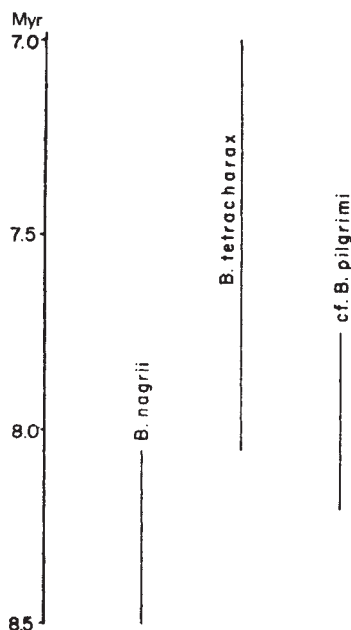


Fig. 2 Temporal ranges in Pakistan of three species of *Brachyrhizomys*.

hominoid level⁶ offer exceptionally precise temporal control on Siwalik fossils in this interval.

Temporal ranges of rhizomyid species of the Siwaliks are well documented and evolutionary patterns of the family are clear³. The primitive genus *Kanisamys* appeared by 13 Myr and represents a long lineage of small rhizomyids that is not presently known from Lufeng. By 9.5 Myr *Brachyrhizomys* evolved, radiating after 8.5 Myr with as many as four large, burrowing contemporaneous species. After 7 Myr both genera became extinct locally, while three new genera appeared. By the Pleistocene, Rhizomyidae disappeared from Pakistan, but they persist in south-east Asia and in Africa. Today Asian rhizomyids occur throughout southern China, south to Sumatra. Most species have wide geographical ranges, for example *Rhizomys pruinosus* (south-east China, west to Assam), which certainly spans >3,000 km of territory.

The presence of at least three advanced species of *Brachyrhizomys* in the Lufeng fauna demonstrates that Lufeng vertebrates cannot be older than 8.5 Myr on the basis of Siwalik rhizomyid phylogeny. Lufeng *Brachyrhizomys* represent species that occur in the Siwaliks and are described below from casts of four specimens, each with three molars. The original material, collected by one of us (Q.G.) is in Beijing.

The smallest specimen (Fig. 1a) is a dentary referable to *B. nagrii*. Size (M_{1-3} length = 11.75 mm) agrees well with Siwalik specimens and M_3 is short (width = 90% of length), a derived character known only in *B. nagrii*. Furthermore, M_2 bears an elongated mesolophid, a primitive trait retained in this species.

Lower and upper dentitions (Fig. 1b, c) are referable to *B. tetracharax*. Although M_2 is moderately worn, the hypoconid has a weak anterior arm. Suppression of this connection is a derived trait of *B. tetracharax* and later rhizomyids. In addition, M_3 is elongated (width = 70% of length) and M_2 bears a rather long mesolophid, which are primitive characters retained in this species. Tooth row length is comparable to Siwalik *B. tetracharax* for both lower (M_{1-3} = 14.88) and upper (M_{1-3} = 12.56) dentitions. The upper row in Rhizomyidae is usually shorter than the lower. Aside from this, identification of this upper dentition is preliminary.

A fourth specimen is bigger than the others (M_{1-3} = 15.6) but smaller than the holotype of *B. pilgrimi*, the largest known species of *Brachyrhizomys*. Although the cast does not reveal diagnostic details of occlusal surfaces, this individual is closest to a population from Pakistan that has been provisionally named³ *Brachyrhizomys* cf. *B. pilgrimi*.

Brachyrhizomys nagrii, *B. tetracharax*, and cf. *B. pilgrimi* have well known temporal ranges in Pakistan and coexisted for a short time at ~8 Myr (Fig. 2). Assuming that rhizomyid species ranged geographically about as far as they do now and had similar temporal ranges in southern China and Pakistan, this places the Lufeng hominoid fauna at ~8 Myr also. Further collecting at Lufeng will improve samples of fossil species and screenwashing will probably add smaller taxa to the faunal list that will help to refine this correlation.

It is probable that hominoids in Pakistan and southern China coexisted in the same broad biogeographical province at 8 Myr. Knowledge of temporal relations of samples of fossil hominoids is essential to reconstructing their pattern of evolution and palaeobiogeography and analysis of other faunal elements can elucidate the palaeoecological setting in which hominoids evolved. Continued synthesis of all available data will offer the fullest understanding attainable of human evolution in southern Asia.

We thank Drs E. H. Lindsay, D. R. Pilbeam and L. L. Jacobs for reviewing the manuscript, and Dr Li Chuan-Kuei for many valuable discussions. We are grateful for the cooperation of S. M. Ibrahim Shah (Geological Survey of Pakistan) and the Institute of Vertebrate Paleontology and Paleoanthropology. The work was partially supported by NSF grants BNS 772-5984 and DEB 7809539 and by SFCP grant FC80254100.

Received 29 March; accepted 28 June 1982.

1. Pilbeam, D. R., Rose, M. D., Badgley, C. & Lipschutz, B. *Postilla* **181**, 1-94 (1980).
2. Kretzoi, M. *Nature* **257**, 578-581 (1975).
3. Flynn, L. J. *Geobios* **15** (in the press).
4. Qi, G. *Vert. Palasiatica* **14**, 14-22 (1979).
5. Chiu, C.-S., Li, C.-K. & Chiu, C.-T. *Ann. Géol. Pays Hellén.* **1**, 263-272 (1979).
6. Tauxe, L. & Opdyke, N. D. *Paleogeogr. Paleoclimatol. Paleocool.* **37**, 43-61 (1982).

Natural selection on body size and proportions in house sparrows

Robert C. Fleischer & Richard F. Johnston

Museum of Natural History and Department of Systematics and Ecology, University of Kansas, Lawrence, Kansas 66045, USA

Variation in morphology across a species range is largely a function of climatic conditions. Evidence for this comes primarily from correlations of latitudinal trends of climate with those of body size or shape (the classic ecogeographic rules of Bergmann and Allen¹). Few studies have documented a temporal sequence of natural selection on size variables within a single population²⁻⁷, and no study precisely documents selection on shape variables. For birds in temperate regions, winter climate can be a major source of mortality⁸, and is likely to be an important agent of selection on overall size²⁻⁶. We reported previously⁶ that severe winter weather produced significant changes in the distribution of size classes of a population of house sparrows (*Passer domesticus*). In that study, post-winter males were on average significantly larger than pre-winter males and post-winter females averaged significantly smaller than pre-winter females. We report here further analyses indicating that sparrows also undergo selection for refinement of bodily proportions, consistent with predictions from Allen's ecogeographic rule. Such selection for increased core-to-limb proportions probably results in greater thermoregulatory efficiency. The changes in proportions accompany the changes in overall size⁶ and there are major inter-sexual differences in how such proportions are achieved.

The study is based on 14 skeletal measurements⁹ from 240 autumn (November 1978) and 196 Spring (March 1979) specimens of house sparrows from Lawrence, Kansas. The 1978-79 winter in Lawrence was one of the most severe on record¹¹. Monthly mean temperatures for December-February averaged 4.8 °C below normal, and snow cover lasted for 55 days of the

Table 1 Selection intensities for skeletal elements of autumn and spring groups

	First-year autumn versus all spring				
	Male (n = 81 vs. 115)			Female (n = 60 vs. 81)	
	Mean (s.d.)		P*	Mean (s.d.)	P*
Skull elements (5)	0.040 +		<0.05	0.008 +	NS
Core elements (3)	0.125 +		<0.01, <0.05	0.042 -	NS
Limb elements (6)	0.059 +		<0.05, <0.05	0.135 -	<0.01, <0.05, <0.05
PC1	0.154 +		<0.01	0.056 -	NS
PC3	0.021 +		NS	0.046 +	NS
	First-year autumn versus first-year spring				
	Male (n = 81 vs. 15)			Female (n = 60 vs. 10)	
	Mean (s.d.)		P*	Mean (s.d.)	P*
Skull elements (5)	0.003 ±		NS	0.030 -	NS
Core elements (3)	0.106 +		<0.05	0.006 ±	NS
Limb elements (6)	0.010 ±		NS	0.105 -	NS
PC1	0.045 +		NS	0.119 -	NS
PC3	0.037 +		NS	0.135 +	NS

Values given are mean selection intensities for element groups (number of elements per group in parentheses), selection intensities for PC scores, and a summary of the significant overwinter differences (by *t*-test) among elements within their respective groupings. + Indicates an increase in the value(s) of the element or PC overwinter, - means a decrease, and ± indicates ambiguity in the direction of selection.

* Significance levels for those elements were significantly different by *t*-test between autumn and spring samples. No variances differed between autumn and spring samples. Univariate analysis presented elsewhere^{6,14}. NS, not significant.

3-month period, compared with about 15 days in an average year.

Within each sex, different comparisons of autumn and spring groups were possible, but only autumn specimens could be satisfactorily identified as first-year birds by the incomplete ossification of their skulls¹². However, a large number of banded specimens was included in the sample, from a previous study at the site¹³. From these birds of known age, samples of first-year spring specimens were identified and compared with first-year autumn birds. These provide the most useful comparison, in terms of measuring levels of wintertime selection, but a comparison is also made of first-year autumn versus all spring specimens. This mixes a number of separate age classes in the spring specimens, but all members of the group are similar in having had at least one wintertime selective experience. First-year autumn birds had completed skeletal growth⁶.

Univariate analyses of the 14 variables over these specimen sets are provided elsewhere^{6,14}. Spring means differed significantly by *t*-test from autumn means for several variables. Variances did not differ among the samples by Levene tests. To analyse the major trends of variation in the skeletal variables, the variance-covariance matrix of the entire log-transformed data set was subjected to principal component analysis¹⁵ (PCA) and PC scores were computed for each specimen. Three principal axes contained significant proportions of variance and were easily interpretable. PC1 contained 42% of the variance and, because of its large, positive loadings for all variables, represents overall size. PC2 (13%) seems to represent a skull to body proportion, and is not included in the present analysis. PC3 contained 10% of the variance, had positive loadings on core variables, negative loadings on limb variables and insignificant loadings on skull variables. Thus, PC3 represents a core-to-limb ratio. Scores on PC1 and PC3 were averaged over each group as defined above.

Here we compare morphology of autumn and spring groups using calculations of selection intensity¹⁶ (measured as the proportional change in mean fitness, $\Delta\bar{w}/\bar{w}$). Selection intensities

estimate the relative amount of morphological change from a pre-selection (autumn) to a post-selection (spring) sample. They are advantageous in that they are standardized and comparable over characters. Because selection acted primarily on the character means in the comparisons, we assume a linear fitness function^{7,16-18}.

Selection intensities were calculated for each of the 14 skeletal elements, grouped according to their respective bodily component (skull, core or limb) and averaged. These mean values, and the selection intensities for the PC scores, are presented in Table 1, which also includes a summary of significant differences. Selection intensities were large for certain characters. For example, $\Delta\bar{w}/\bar{w}$ was 0.237 for decrease in humerus length in females. By comparison, $\Delta\bar{w}/\bar{w}$ for change in humerus length from O'Donald's¹⁷ re-analysis of Bumpus' data on storm-generated selection in *P. domesticus* was only 0.103. Drought-mediated selection on *Geospiza fortis*⁷ produced the largest intensities ever noted: for example, $\Delta\bar{w}/\bar{w}$ on wing chord over a 2-yr period was 0.39.

The pattern of selection intensities differs to some extent between the sexes. On PC1, males show strong selection for increased body size, but females show strong selection for decreased body size. Selection intensities on PC3 are weak to moderate; mean changes are not significant, but all indicate an increased core-to-limb ratio. Selection intensities averaged over body components indicate more clearly a reportioning of body parts. In males, selection intensities over skull and limb elements are weak, although generally showing a trend towards increased size. In comparison, average selection intensities for core elements are considerably larger for both the first-year autumn versus all spring comparison and for the first-year autumn versus first-year spring comparison. Male core elements, therefore, show a disproportionate increase relative to their limb or skull elements.

For females, the basic pattern of selection is opposite to that of males. Females exhibit low selection intensities in variable directions for skull and core characters, but extremely strong selection for decreased limb size. Females limb elements thus show a disproportionately greater decrease relative to their core or skull elements. In summary, Table 1 reveals that while spring males are larger overall and spring females are smaller overall, each sex has also experienced a differing mode of selection on body components that results for both sexes, in an increased core-to-limb ratio.

Individuals with a high core-to-limb ratio clearly have a selective advantage in severe winter weather. A decreased surface-to-mass ratio results from such a core-to-limb trend, which in turn may result in lower rates of metabolism and heat loss^{19,20}. Such trends in the proportioning of body parts are evident in geographical comparisons within a number of species¹, and within house sparrows such a cline also exists⁹. However, we believe this report to be the first to indicate selection for proportioning of body parts on a temporal basis. As there are constraints on body size in each sex, for males to be larger and females smaller post-selection, there must also be constraints on the way each sex accommodates selection for optimal proportioning of body parts. Spring males must have greatly increased cores in relation to their also larger limb elements, and spring females must have considerably smaller limbs in relation to their also smaller core elements.

In house sparrows, therefore, males and females are selectively distinct morphometric entities: sexual selection results in larger-sized males²¹; and females perhaps undergo selection for small size because of reproductive constraints²². As shown here, given intense selection pressures, overwinter mortality can support or augment these trends⁶. Even so, males and females, although probably allowed to differ in size because of sex-limited loci, cannot be totally distinct in an evolutionary sense because of the effects of syngamy and recombination. Sexual reproduction annually generates cohorts of offspring of interparental size or shape that cannot wholly anticipate wintertime climate. Selection sometimes then acts on this new

cohort to generate post-winter phenotypic distributions better approximating optimal distributions.

Support of the NSF (DBS-7912412) and the General Research Fund of the University of Kansas is hereby acknowledged.

Received 3 November 1981; accepted 27 May 1982.

1. Mayr, E. *Evolution* **10**, 105–108 (1956).
2. Bumpus, H. C. *Biol. Lect., Marine Biol. Lab. Woods Hole*, 1–15 (1898).
3. Johnston, R. F., Niles, D. M. & Rowher, S. A. *Evolution* **26**, 20–31 (1972).
4. Rising, J. D. in *Productivity, Population Dynamics and Systematics of Granivorous Birds* (eds Kendeigh, S. C. & Pinowski, J.) (Polish Scientific Publishers, Polish Academy of Sciences, Warszawa, 1972).
5. Baker, M. C. & Fox, S. F. *Am. Nat.* **112**, 675–682 (1978).
6. Johnston, R. F. & Fleischer, R. C. *Auk* **98**, 503–511 (1981).
7. Boag, P. T. & Grant, P. R. *Science* **214**, 82–85 (1981).
8. Fretwell, S. D. *Populations in a Seasonal Environment* (Princeton University Press, 1972).
9. Johnston, R. F. & Selander, R. K. *Evolution* **21**, 1–28 (1971).
10. Johnston, R. F. *Systematic Zool.* **22**, 219–226 (1973).
11. Matson, M. & Weisnet, D. R. *Nature* **289**, 451–453 (1981).
12. Niles, D. M. *Condor* **75**, 354–356 (1973).
13. Lowther, P. E. thesis, Kansas Univ. (1979).
14. Fleischer, R. C. thesis, Kansas Univ. (1982).
15. Harris, R. J. *A Primer of Multivariate Statistics* (Academic, New York, 1975).
16. O'Donald, P. *Theor. Population Biol.* **1**, 219–232 (1970).
17. O'Donald, P. *Evolution* **27**, 398–404 (1973).
18. Lowther, P. E. *Evolution* **31**, 649–656 (1977).
19. Kendeigh, S. C. *Auk* **86**, 13–25 (1969).
20. Blem, C. R. *Comp. Biochem. Physiol.* **47A**, 101–108 (1974).
21. Selander, R. K. in *Sexual Selection and the Descent of Man* (ed. Campbell, B.) Ch. 8 (Aldine, Chicago, 1972).
22. Downhower, J. F. *Nature* **263**, 558–563 (1976).

Genetic differences in individual behaviour associated with shell polymorphism in the snail *Cepaea nemoralis*

J. S. Jones

Department of Genetics and Biometry, University College London, Wolfson House, 4 Stephenson Way, London NW1 2HE, UK

It is often supposed that genetic polymorphism in a population will be favoured if individuals of different genotype occupy an ecological niche in which they are relatively fit¹. There are great practical difficulties in establishing whether such genetic differences in behaviour influence the niches chosen by polymorphic animals in nature. I report here a new technique for investigating habitat selection in the land snail *Cepaea nemoralis*. This species is polymorphic for shell colour and banding and hence for shell reflectivity. Differences in the efficiency of absorption of solar energy by the various phenotypes lead to geographical changes in allele frequency². In this study snails of different phenotype in the same population were marked with a paint which fades at a measurable rate when exposed to daylight. This detects individual differences in daytime activity too slight to be identified by direct observation and shows that within a population there are differences in the behaviour of individuals of contrasting shell phenotype when exposed to sunlight. Habitat selection may therefore be partially responsible for the extensive genetic polymorphism of *C. nemoralis*.

There are great differences in gene frequency among geographically separated populations of *C. nemoralis*. Some of these differences result from climatic selection; dark-coloured snails absorb more solar energy than do light-coloured, and hence may be at an advantage in cold and shady microclimates but liable to death through overheating in warm sunny places^{2–4}. Much less is known of the forces which maintain genetic variation within a population than of those which lead to genetic change among populations. Individual differences in thermal relations in sunshine might influence the habitat chosen by snails of different genotype, so that the shell polymorphism may be associated with behavioural differences which help to maintain the various alleles in the population.

Observation of differences in behaviour in natural populations is difficult as the snails are often buried or hidden by vegetation. What is needed is to measure the activity of single

snails over a period long enough to detect slight individual differences. Most techniques of measuring exposure to daylight are unsuitable as they involve complex instruments or undue physical interference⁵. There are, however, several chemical or electronic methods of integrating solar energy falling on a surface. Chemicals which degrade at a known rate when exposed to the Sun^{6–8} are not usually practical for field use as they involve the spectrophotometry of degradation products. I have modified a technique used to measure the stability of textile dyes to give a convenient method of assessing the behaviour of individual animals in daylight.

The Blue Scale is a series of pigments of differing resistance to fading which are dyed onto cloth⁹. The technique can be improved for field use by dissolving the pigments in a stable yellow paint to give a green paint which fades to yellow. The extent of fading is assessed by comparison with a scale produced by mixing this paint with increasing proportions of the yellow base. The paint is primarily sensitive to UV light (which represents only a small part of the Sun's thermal energy¹⁰) so that paint fading integrates exposure to daylight over time rather than directly assessing thermal stress. Paint fading gives an indication of the activity of individuals over an extended period.

A saturated solution of benzyl blue in monoethyleneglycol-monoethylether was added to yellow polyurethane paint (International Paint Co., Series 101) to produce a dark green compound. Experimental snails were collected in the valley of the Garonne in the Spanish Pyrenees, where *C. nemoralis* is highly polymorphic¹¹. Behavioural responses were tested in semi-natural conditions in cages in the field at Nettlecombe Court Field Centre, Somerset, UK. These cages have been used by Bantock¹² to study mortality in *C. nemoralis*. His results seem to reflect the fitness differences between shell morphs which exist in natural populations. The cages have the great practical advantage that all surviving marked snails can be recaptured.

Twenty adult yellow unbanded and the same number of yellow five-banded snails were collected from each of 10 natural populations of *C. nemoralis* during the period 19–23 June 1981. In every case snails were collected from an area smaller than that of a panmictic population¹³. Each snail was marked with a small spot of paint placed 5 mm behind the lip at the equator of the shell, and the populations were placed in 10 field cages planted with nettles and grass on 3 July 1981. The extent of paint fading was measured 60 days later (Table 1).

There are significant differences between banded and unbanded snails in mean exposure to daylight. In every cage,

Table 1 Degree of paint fading on banded and unbanded snails

Cage		Faded to point							
		1	2	3	4	5	6	7	8
1	U		3	6	7	3			
	B			1	3	10	4		
2	U		3	3	5	3			
	B				6	9	1		
3	U		2	10	3	3			
	B			6	8	1			
4	U			4	7	4	3		
	B			1	3	12	3		
5	U		1	6	7	3			
	B			1	11	6	1		
6	U			4	9	7			
	B			4	3	10	3		
7	U		2	10	5	1			2
	B			2	7	9	1		
8	U			8	10				
	B				14	4			
9	U			1	9	8	2		
	B				4	12	2		1
10	U			7	9	2	1		
	B				11	6			

Paint fading on green to yellow scale of marked *C. nemoralis* in field population cages, U, unbanded snails; B, banded.

unbanded shells show a lower mean paint fading than do banded. This is in itself significant (sign test: $P = 0.002$), and in five cages the effect is shown by a Mann-Whitney U -test to be individually significant at the 5% level. The overall difference between the mean fading of unbanded (3.85 points) and banded (4.55 points) is highly significant ($t = 7.54$; $P < 0.001$). In addition, in 9 of the 10 cages there is a greater variance of fading among unbanded shells (sign test: $P = 0.022$). There are therefore considerable behavioural differences associated with the shell polymorphism of *C. nemoralis*.

Partitioning of the environment by snails of different genotype may result either from temporal or from spatial differences in activity. There is a general tendency for related coexisting species of small poikilotherm to differ in their times of activity¹⁴. *C. nemoralis* and *Cepaea hortensis* differ in their period of maximum activity in a way which is related to their preferred microclimate¹⁵. Australian *Theba pisana* show differences in the seasonal activity of dark and light-coloured snails which depend on their thermal relations in sunshine¹⁶. This effect might also be important in *C. nemoralis*, so that the dark-coloured banded shells are active for a higher proportion of the day in the cool English climate.

Banded and unbanded snails might also choose to live at a different height above the surface. Behaviour in sunlight is the prime mode of thermoregulation in invertebrates¹⁷. Temperature stresses are often greater near the ground because of the development of a layer of heated air on sunny days¹⁰ and many organisms exposed to solar stress choose to live at some distance above the surface. Mediterranean *Helicella virgata*, for example, reach a lethal temperature much more often when resting at 3 cm above the ground than when at 70 cm (ref. 18). In *C. nemoralis*, also, snails climb most actively in hot dry weather¹⁹. Behavioural thermoregulation of this kind is likely to be particularly important in banded *C. nemoralis*, which are more liable than are unbanded to die of heat shock when exposed on the ground surface on sunny days²⁰. Banded snails will therefore climb higher than do unbanded, so that their paint spots will fade to a greater extent. The greater variance of fading scores among unbanded snails suggests that they have a wider range of behaviours and hence a broader niche with respect to general activity, perhaps because they can penetrate both into the layer of heated air near the ground and the direct rays of the Sun at the top of the vegetation. Experiments with *Drosophila* show that temporal fluctuations in habitat are as effective in maintaining polymorphism as are spatial changes²¹, so that differences in either time or location of activity by snails of different shell phenotype might favour the evolution of shell polymorphism.

In laboratory populations of various organisms, genetic differences in habitat selection increase the fitness of particular alleles²²⁻²⁴, and allow a *Drosophila* mutant to be maintained in a heterogeneous environment although it is otherwise harmful to its carriers²⁵. Very few examples of genetic polymorphism for habitat selection have been established in natural or semi-natural conditions^{16,26}. The experiments described here show that different shell genotypes of *C. nemoralis* have differences in patterns of activity which may influence their fitness and hence promote shell polymorphism. This technique of integrating activity over time may be generally useful in exploring the behavioural ecology of the individual differences thought to exist in various animals as they search for food, for mates or for territory.

This work was supported by the NERC and the Royal Society. I thank the Field Studies Council and Dr S. M. Tilling for help at Nettlecombe Court; Dr C. R. Bantock for the use of his field cages; Professor C. A. B. Smith for statistical advice; Messrs S. Borfield, P. Slepokura and D. Young for help in the field; and Dr M. Tordoff of the Society of Dyers and Colourists, Dr A. Davis of the Propellants, Explosives and Rocket Motors Establishment and Dr K. L. Rabone of Unilever Research for advice on measuring solar radiation.

Received 1 April; accepted 24 May 1982.

1. Taylor, C. E. *Genetics* **83**, 887-894 (1976).
2. Jones, J. S., Leith, B. H. & Rawlings, P. A. *Rev. ecol. Systematics* **8**, 109-143 (1977).
3. Cain, A. J. *Phil. Trans. R. Soc. B253*, 499-517 (1968).
4. Jones, J. S. *Science* **182**, 546-552 (1973).
5. Monteith, J. L. *Int. biol. Progr. Hbk* **22**, 1-263 (1972).
6. Davis, A., Deane, G. H. W., Gordon, D., Howell, J. V. & Ledbury, K. J. *J. appl. Polym. Sci.* **20**, 1165-1174 (1976).
7. Davies, R. E. in *The Biological Effects of Ultraviolet Radiation* (ed. Urbach, F.) 437-443 (Pergamon, Oxford, 1969).
8. Diffey, B. L. & Davis, A. *Phys. Med. Biol.* **23**, 318-323 (1978).
9. *British Standard* No. 1006 (British Standards Institution, London, 1961).
10. Monteith, J. L. *Principles of Environmental Physics* (Arnold, London, 1973).
11. Arnold, R. W. *Phil. Trans. R. Soc. B253*, 549-593 (1968).
12. Bantock, C. R. *J. Linn. Soc. Biol.* **13**, 47-64 (1980).
13. Lamotte, M. *Bull. biol. Fr. Belg. Suppl.* **35**, 1-239 (1951).
14. Schoener, T. W. *Science* **185**, 27-39 (1974).
15. Cameron, R. A. D. *J. Zool.* **162**, 303-315 (1970).
16. Johnson, M. S. *Heredity* **47**, 121-134 (1981).
17. Henwood, K. *Ecology* **56**, 1329-1342 (1975).
18. Pomeroy, D. E. *Aust. J. Zool.* **16**, 857-869 (1968).
19. Jaremovic, R. & Rollo, C. D. *Can. J. Zool.* **57**, 1010-1014 (1979).
20. Richardson, A. M. M. *Nature* **247**, 572-573 (1974).
21. Mackay, T. F. C. *Genet. Res.* **37**, 79-94 (1981).
22. Christensen, B. *Heredity* **87**, 21-26 (1977).
23. Cavener, D. *Behav. Genet.* **9**, 359-365 (1979).
24. Steward, R. C. *J. Zool.* **178**, 107-115 (1976).
25. Jones, J. S. & Probert, R. F. *Nature* **287**, 632-633 (1980).
26. Schoener, T. W. & Schoener, A. *Evolution* **30**, 650-658 (1976).

Pluripotent embryonic stem cell lines can be derived from t^{w5}/t^{w5} blastocysts

Terry Magnuson*, Charles J. Epstein†‡, Lee M. Silver‡ & Gail R. Martin§

Departments of * Pediatrics, † Biochemistry and Biophysics and § Anatomy, University of California, San Francisco, California 94143, USA

‡ Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

Mouse embryos homozygous for t^{w5} , a recessive lethal mutation in the t complex located on chromosome 17, develop normally until the elongated egg cylinder stage, approximately 6.5 days after fertilization. At this time, the endoderm is morphologically abnormal and the embryonic ectoderm begins to show signs of pyknosis. Death of the embryo usually occurs within the next 2 days (ref. 1). A serious difficulty in the study of lethal t -mutant gene expression during embryogenesis has been to obtain appropriate experimental material, particularly during the period immediately following implantation. Recently, a method involving the use of teratocarcinoma-conditioned medium was devised for establishing pluripotent stem cell cultures directly from the inner cell mass (ICM) of a normal mouse blastocyst². Using this method, we have established three embryonic stem cell lines derived from embryos carrying t^{w5} . We report here that one of the cell lines is homozygous for the mutation (t^{w5}/t^{w5}), whereas the other two are heterozygous ($+/t^{w5}$). When injected into athymic mice, each cell line is capable of forming tumours that contain differentiated derivatives of all three primary germ layers.

The three embryonic stem cell lines were obtained from cultures of 10 ICMs isolated from t^{w5} -bearing blastocysts. Mass cultures of the cell lines are indistinguishable from other established embryonic stem cell and teratocarcinoma stem cell lines in both morphological and growth characteristics. When tested for expression of the SSEA-1 cell-surface antigen³, which is common to teratocarcinoma stem cells and early embryos but not to most differentiated cell types, all three cell lines derived from t^{w5} -bearing embryos were found to be positive by indirect immunofluorescence.

It would be desirable, of course, to identify the genotypes of these embryonic stem cell lines by determining whether they produce the wild-type or mutant t^{w5} gene products. However, as these products are unknown, we have used other genetic markers which identify three different regions of the t^{w5} -bearing chromosome. The validity of the markers depends on the fact that the t^{w5} lethal mutation is carried in the context of a

complete t^{w5} haplotype. This t^{w5} haplotype, like all other complete t haplotypes, represents a variant form of a large region of chromosome 17 extending from the T locus near the centromere to beyond the $H-2$ complex (reviewed in ref. 4). In addition to the mutation affecting embryonic development, which for t^{w5} has been mapped near the centre of the complete t^{w5} haplotype⁵, all t haplotypes carry variant alleles of several other genes, including factors that affect sperm differentiation and function. The integrity of a t haplotype is maintained by an extreme suppression of meiotic recombination between t -haplotype and wild-type chromatin (for reviews see refs 4,6-8).

The first marker we used identifies the presence or absence of a t^{w5} -bearing chromosome. The embryos from which the stem cell lines were derived were obtained by crossing

$t^{w5}/Rb(16.17)7Bnr$ mice. In these embryos, the t^{w5} mutation is carried on an acrocentric chromosome 17, whereas the wild-type allele is carried on the $Rb(16.17)7Bnr$ translocation chromosome (comprising chromosomes 16 and 17 fused at the centromere) which thus serves as a marker for its presence. Consequently, in the absence of recombination, chromosome spreads from embryonic stem cells homozygous for t^{w5} will not have a translocation chromosome, heterozygous cells will have one and wild-type cells will have two. When the karyotype of embryonic stem cell line 1 was examined, none of the spreads contained a translocation chromosome (Fig. 1a). In contrast, embryonic stem cell lines 2 and 3 both had a translocation chromosome (Fig. 1c).

Because there is a 4-6% chance of meiotic recombination occurring between the centromere and the start of the t -

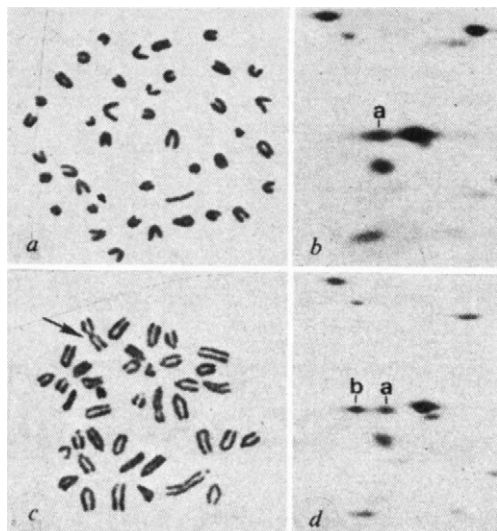


Fig. 1 Identification of embryonic stem cell genotype. Ten ICMs were isolated from blastocysts obtained by mating $t^{w5}/Rb(16.17)7Bnr$ mice. (As the transmission ratio of the t^{w5} haplotype on our background is 87%, we would expect four to five of the ICMs to be t^{w5} homozygotes.) They were seeded on confluent monolayers of mitomycin C-treated fibroblasts in Dulbecco's modified Eagle's medium supplemented with 10% calf serum, 10^{-4} M 2-mercaptoethanol and concentrated conditioned medium². Within 1 week, three of the ICMs had proliferated to form undifferentiated stem cell colonies. These colonies were disaggregated by treatment with trypsin and, with the use of a micropipette, were subcultured to new feeder layers. The cell populations were then expanded according to the procedures outlined by Martin², except that the addition of concentrated conditioned medium was discontinued after three passages. When the karyotype of embryonic stem cell line 1 was examined (a), 38% of the cells had 40 chromosomes, 30% had 41 and 10% had 42. A translocation chromosome was not observed in any spread. In addition, only the p63a protein was found marked a in (b). These data identify line 1 as homozygous for t^{w5} . 39% of the cells of embryo-derived cell line 2 were found to have 42 chromosomes and 22% had 41. 84% of the cells of line 3 had 40 chromosomes. The presence of a translocation chromosome (arrowed in (c) and p63a and p63b (a and b respectively in (d) identifies lines 2 and 3 as heterozygous for t^{w5} . For karyotyping, 10^7 stem cells were seeded in 60 mm² tissue culture dishes and cultured overnight. They were then incubated for 30 min in medium containing $10 \mu\text{g ml}^{-1}$ colcemid (Gibco), washed with phosphate-buffered saline (PBS) and collected after exposure to trypsin. The cells were centrifuged, and the pellet resuspended in 10 ml of 0.56% KCl. After 25 min, the cells were again centrifuged and the resulting pellet resuspended in 10 ml of 3:1 methanol/acetic acid fixative. The fixative solution was changed three times and then refrigerated overnight. The following day, the pellet was resuspended in 1 ml of a freshly made fixative solution, and chromosome spreads were prepared by placing a drop of the cell suspension on cleaned, wetted microscope slides. The chromosomes were stained with 3% Giemsa (GURR) for 5 min. For two-dimensional gel electrophoresis, embryonic stem cells were separated from feeder layer cells by a double preplating technique². The stem cells were then seeded in 16 mm² multi-well tissue culture plates (Flow Labs) at a density of 10^6 and cultured overnight. After washing with PBS, the stem cells were incubated in methionine-free medium for 30 min and then for 5 h in methionine-free medium containing $250 \mu\text{Ci ml}^{-1}$ of ^{35}S -methionine (specific activity of 900-1,100 Ci mmol⁻¹; Amersham). The labelled cells were washed with PBS and lysed with 100 μl of isoelectric focusing sample buffer¹⁸. Two-dimensional gel electrophoresis was performed as outlined by O'Farrell¹⁸ with a 4.5% stacking gel and a 10% separating gel being used in the second dimension. Kodak NS-2T film was used for autoradiography of the dried gels. Only the area of the gel pattern containing p63 is shown.

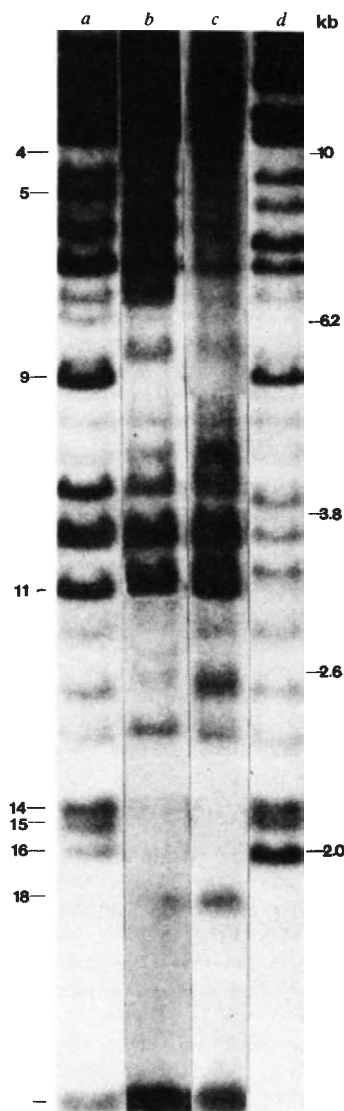


Fig. 2 Blot hybridization analysis of $H-2$ -like genes from embryonic stem cell line 1. *EcoRI* restriction patterns of a, $129(t^{w5}/t^{w5})$; b, $129(+/t^{w5})$; c, embryonic stem cell line 1; and d, $Rb(16.17)7Bnr/Rb(16.17)7Bnr$. High molecular weight DNA was prepared from adult livers or embryonic stem cells as described previously¹³. DNA was digested to completion with *EcoRI* and 5- μg aliquots of each sample were electrophoresed through a 1% agarose gel at 0.7 V cm^{-1} for 36 h. Southern blot hybridization was performed with clone pHIIa DNA¹⁹ nick-translated to a specific activity of $2 \times 10^8 \text{ c.p.m. } \mu\text{g}^{-1}$, and used at $8.2 \times 10^4 \text{ c.p.m. cm}^{-2}$ of filter. Hybridization and washing conditions were as described previously. Particular t -specific *EcoRI* fragments are numbered at the left according to Silver¹³. The numbers at the right denote the size and positions of the DNA length standards, in kilobases (kb). Lanes b and c are composites of different exposures of the same gel.

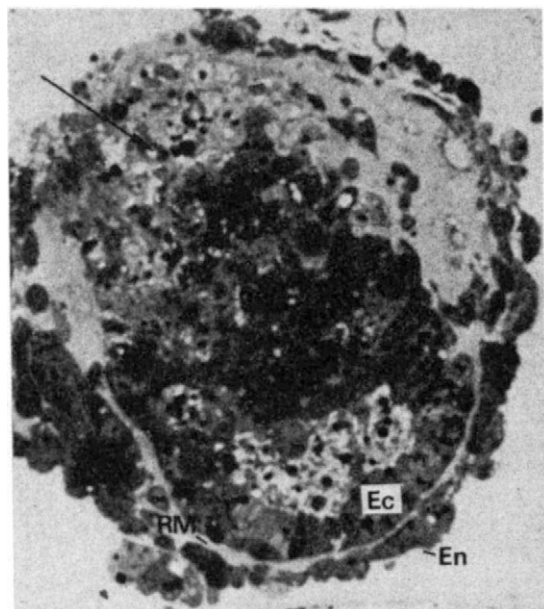


Fig. 3 Section of an embryoid body formed by t^{w5}/t^{w5} embryonic stem cells. Aggregates of cells were cultured in Petri dishes for 16 days. The embryoid bodies which formed were fixed in 2.5% glutaraldehyde, dehydrated and embedded in Epon, and 1.5- μ m sections were prepared. Part of a cavitating embryoid body, consisting of columnar ectoderm (Ec) surrounded by a layer of endoderm-like cells (En) with underlying Reichert's membrane-like material (RM), is shown. There are numerous pyknotic cells (arrow) in this embryoid body. $\times 240$.

haplotype region⁹, the second marker we used was the *Tcp-1* gene located in the region of the t^{w5} haplotype proximal to the lethal gene¹⁰. All complete t haplotypes, including t^{w5} , carry a variant *Tcp-1* allele (*Tcp-1*^a) which encodes a polypeptide p63/6.9a (abbreviated p63a) that is readily distinguishable from the wild-type gene (*Tcp-1*^b) product (p63/6.9b or p63b) by two-dimensional gel electrophoresis¹¹. Thus, t^{w5} homozygous cells express only the p63a form of the polypeptide, heterozygous cells express both p63a and p63b in equal amounts and wild-type cells express only the p63b form. When the protein gel pattern of embryonic stem cell line 1 was examined, only the p63a form of the polypeptide was observed (Fig. 1b). In contrast, both p63a and p63b were expressed in roughly equal amounts by embryonic cell lines 2 and 3 (Fig. 1d).

As meiotic recombination within the t complex is highly suppressed (recombination frequency between 0.001 and 0.002)⁶, the probability that *Tcp-1* has been separated from the t^{w5} lethal mutation and that cell line 1 is not homozygous for t^{w5} is extremely low (≤ 0.004). Nevertheless, we used the family of genes comprising the *H-2* complex as a third marker to demonstrate that cell line 1 carries only the t^{w5} haplotype. The *H-2* complex is located in the region of the t^{w5} haplotype distal to the lethal factor⁵. t -haplotype-specific restriction fragments carrying DNA sequences homologous to the *H-2* class I genes have recently been described^{12,13}. With the use of the restriction enzyme *EcoRI*, multiple t -haplotype-specific fragments can be observed in the context of the 129/SvJ background pattern (Fig. 2, lane a). Although fragments 5, 11 and 17 are shared by all complete t haplotypes, fragments 4 and 18 are unique to the t^{w5} haplotype (Fig. 2, lane b) and are not present in genomic DNA from other lethal and semilethal haplotypes such as t^0 , t^{w2} , t^{w12} or t^{12} . All the other well characterized t haplotypes display fragments 9, 14, 15 and 16 whereas t^{w5} does not¹³. The *H-2* haplotype associated with the Rb(16.17)7Bnr chromosome also displays *EcoRI* fragments which migrate in the same regions as t fragments 9, 14–16 (Fig. 2, lane d).

EcoRI-digested DNA from embryonic cell line 1 displays t -specific *H-2* fragments 5, 11 and 17, and t^{w5} -specific frag-

ments 4 and 18 (Fig. 2, lane c). This pattern of *EcoRI* fragments is characteristic of the t^{w5} haplotype. The absence of fragments 9, 14–16 rules out the unlikely possibility that a lethal or semi-lethal haplotype other than t^{w5} is present in this cell line, and also verifies the absence of the *H-2* genomic region of the Rb(16.17)7Bnr chromosome.

Taken together, the analyses of the three markers demonstrates conclusively that line 1 is homozygous for regions of the t^{w5} haplotype both proximal and distal to the t^{w5} lethal gene. As suppression of recombination along the t^{w5} haplotype reduces the probability of a double recombination event to $\sim 10^{-6}$, we can definitely conclude that this cell line is homozygous for the t^{w5} lethal gene.

To determine whether the embryonic stem cells can mimic early embryogenesis, the homozygous and one of the hetero-

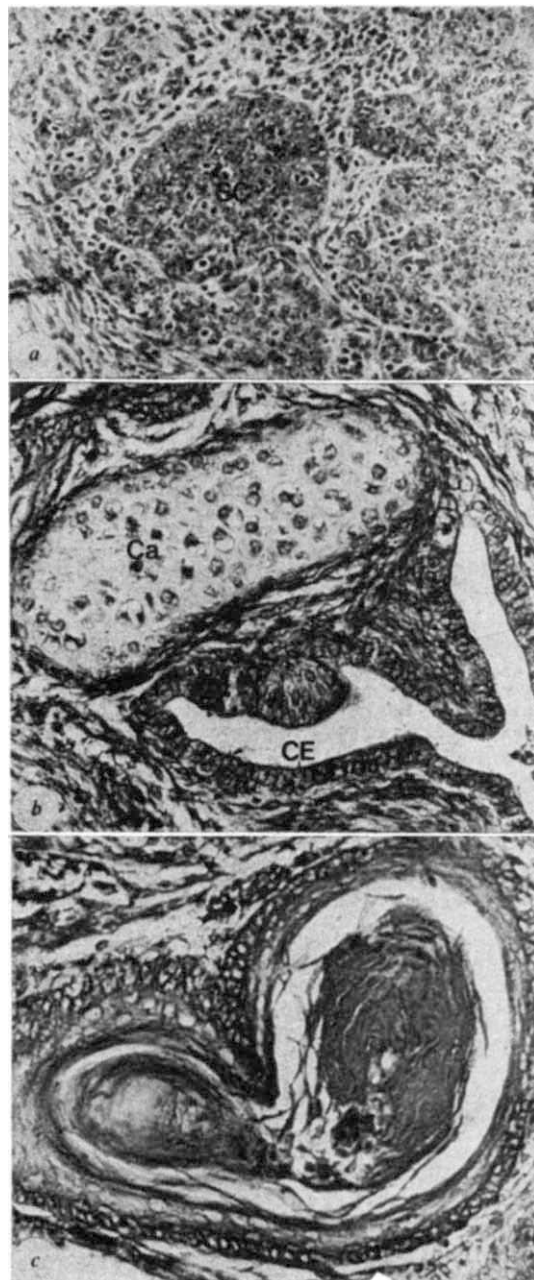


Fig. 4 Sections of tumours formed after subcutaneous injection of 5×10^6 t^{w5}/t^{w5} embryonic stem cells into an athymic nude mouse. After 2–6 weeks the tumours that had formed were removed, fixed in Bouin's solution, embedded in paraffin and sectioned. The cell types observed in the tumours included a, groups of undifferentiated stem cells (SC); b, cartilage (Ca) and columnar epithelium (CE); and c, keratinizing epithelium. a, $\times 80$; b, $\times 104$; c, $\times 112$.

zygous cell lines were cultured in appropriate conditions¹⁴. Both cell lines were able to form embryoid bodies. Under the light microscope, the t^{w5}/t^{w5} embryoid bodies were observed to consist of a core of ectoderm-like cells surrounded by a layer of endoderm-like cells. Typical Reichert's membrane-like material, presumably secreted by the endodermal cells, was observed in almost all the t^{w5} -homozygous embryoid bodies, and the beginning of pro-amniotic cavity formation was observed in the core of some (Fig. 3). Extensive death of core cells was observed in some of the t^{w5}/t^{w5} embryoid bodies. However, although ectoderm degeneration is characteristic of t^{w5}/t^{w5} embryos *in utero*, the cell death observed in the t^{w5}/t^{w5} embryoid bodies cannot necessarily be attributed to the effects of the t^{w5} lethal mutation as numerous pyknotic cells have also been observed in the core of embryoid bodies formed by some wild-type teratocarcinoma stem cell lines (G.R.M., unpublished data).

As the variety of cell types formed during the development of embryoid bodies is always limited, each of the t^{w5} -bearing embryonic stem cell lines was injected into several athymic mice (*nu/nu*) to determine the extent of their differentiative capacities. Within 2–6 weeks, all the cell lines produced tumours containing undifferentiated stem cells as well as differentiated derivatives of all three primary germ layers. The differentiated cell types included neural, keratinizing, pigmented and columnar epithelium, yolk sac endoderm, cartilage, muscle and connective tissue (Fig. 4a–c). Similar results were obtained with a subclone of the t^{w5}/t^{w5} embryonic stem cell line. We thus conclude that the t^{w5} -homozygous embryonic stem cell line is pluripotent and capable of differentiating into a large number of differentiated cell types. Many of these cell types are not formed by t^{w5}/t^{w5} embryos *in utero* because the embryo dies before all three primary germ layers are established¹.

The mechanisms of *t*-mutant lethality are not yet understood. One hypothesis is that at least the earlier-acting lethal *t* mutations, such as t^{12} , t^0 and t^{w5} , are generalized cell lethals that cause embryonic death by perturbing processes such as intermediary metabolism, which are essential for the survival of all cells in the embryo¹⁵. Wudl and Sherman¹⁶ interpreted their observation that all cells in cultured presumptive t^{w5}/t^{w5} blastocysts die at earlier times than do cells of control embryos as evidence that t^{w5} is a generalized cell lethal. Hogan *et al.*¹⁷ also showed that presumptive t^{w5}/t^{w5} ICMs are smaller, abnormally organized and apparently moribund when compared with ICMs from normal embryos. Our results, however, demonstrate that t^{w5}/t^{w5} ICM cells are sufficiently viable to give rise to an embryonic stem cell line which, in the appropriate environment, can differentiate into a wide variety of cell types. The simplest explanation of these results is that the t^{w5} mutation does not result in generalized cell lethality. Death of t^{w5}/t^{w5} embryos must therefore be due to other causes. The availability of the t^{w5} -homozygous embryonic stem cell line described here should make possible new approaches to the study of how such lethal *t* mutations affect normal development.

We thank Estrella Lamella, Theodosia Zamora, Marianne Gallup and Jane Uman for technical assistance and Dave Akers for assistance in preparing the figures. This work was supported by NIH grants HD-03132, HD-16091, CA-25966 and basic research grant 1-770 from the March of Dimes Birth Defects Foundation. T.M. is a recipient of a NIH National Research Service Award and G.R.M. of an American Cancer Society Faculty Research Award.

Received 6 May; accepted 15 June 1982.

- Bennett, D. & Dunn, L. C. *J. Morph.* **103**, 135–157 (1958).
- Martin, G. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7634–7638 (1981).
- Solter, D. & Knowles, B. B. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5565–5569 (1978).
- Silver, L. M. *Cell* **27**, 239–240 (1981).
- Artzt, K., Shin, H.-S. & Bennett, D. *Cell* **28**, 471–476 (1982).
- Bennett, D. *Cell* **6**, 441–454 (1975).
- Sherman, M. I. & Wudl, L. R. in *Concepts in Mammalian Embryogenesis* (ed. Sherman, M. I.) 136–234 (MIT Press, 1977).
- Lyon, M. F. *Symp. zool. Soc. Lond.* **47**, 455–477 (1981).
- Magnuson, T. & Epstein, C. J. *Biol. Rev.* **56**, 369–408 (1981).
- Silver, L. M., White, M. & Artzt, K. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6077–6080 (1980).

- Silver, L. M., Artzt, K. & Bennett, D. *Cell* **17**, 275–284 (1979).
- Shin, H.-S., Stavenezer, J., Artzt, K. & Bennett, D. *Cell* (in the press).
- Silver, L. M. *Cell* (in the press).
- Martin, G. R., Wiley, L. M. & Damjanov, I. *Devl Biol.* **61**, 230–244 (1977).
- Wudl, L. R., Sherman, M. I. & Hillman, N. *Nature* **270**, 137–140 (1977).
- Wudl, L. R. & Sherman, M. I. *Cell* **9**, 523–531 (1976).
- Hogan, B., Spiegelman, M. & Bennett, D. *J. Embryol. exp. Morph.* **60**, 419–428 (1980).
- O'Farrell, P. H. *J. biol. Chem.* **250**, 4007–4021 (1975).
- Steinmetz, M. *et al. Cell* **24**, 125–134 (1981).

Plasminogen activator–plasmin system and neuronal migration

G. Moonen*, M. P. Grau-Wagemans† & I. Selak*

* Département de Clinique et de Pathologie médicales, Secteur de Neurologie and † Service de Cytologie et d'Histologie, Université de Liège, 20 rue de Pitteurs, 4020 Liège, Belgium

Neurogenesis results from a synchronized series of elementary events including cellular proliferation, migration, differentiation, recognition and death. Neuronal migration is a key step in neural morphogenesis since inadequately located neurones may not establish the appropriate connections and this may lead to neuronal death or to functional deficit of synaptic circuits. Impairment of neuronal migration has been implicated in human pathology¹ and well documented in animal pathology, such as the *weaver* mutation in mice². The cerebellum of small rodents is particularly well suited for the study of neuronal migration because a subpial neuronogenesis occurs postnatally. The subsequent inward migration of postmitotic neurones (mostly granule cells) has been studied using classical neuroanatomical methods, such as the Golgi stain³, and autoradiography after systemic injection of ³H-thymidine⁴. Systematic ultrastructural investigation of neuronal migration at different levels of the central nervous system and in different species led Rakic to propose the radial glia hypothesis—that neurones migrate along radial glia cells which serve as guides during migration⁵. The observation that neurones migrate inside a densely packed neuropile has prompted us to consider the possible role of extracellular neutral proteolysis during neuronal migration. We have focused on the plasminogen activator (PA) serine proteases because these enzymes are known to be involved in several phenomena that involve cell migration or tissue remodelling⁶. The usual substrate for PA is plasminogen, which is converted to plasmin, although other substrates may exist⁷. We report here that both PA and plasmin are released by cultured 7-day-old rat paraflocculus, but not by 1-month-old or adult rat paraflocculus (that is, after granule cell migration), and that inward migration of cerebellar granule neurones, which account for 95% of the cells in adult cerebellum, can be inhibited by inhibitors of the PA–plasmin system.

We have studied PA and plasmin activity with the use of the classical fibrin plate assay^{8,9}. When standard plates are used, it is possible to demonstrate both endogenous PA and plasmin activity; with heated plates only endogenous plasmin activity can be detected (see Fig. 1 legend). When a 7-day rat cerebellum paraflocculus (the most lateral part of the cerebellum) was explanted for 24 h at 37°C on a standard fibrin plate, a fibrinolytic zone formed around it; the fibrinolysis could be inhibited by 6-aminocaproic acid. In the same conditions, but using heated fibrin plates, a smaller but definite fibrinolytic zone could be seen, demonstrating that not only PA but also plasmin are released by the 7-day rat paraflocculus. If these enzymes were to have a role in neuronal migration, we reasoned that 1- or 3-month (adult) rat paraflocculus should not release PA and plasmin. We found that, indeed, provided the meninges were carefully dissected, in neither case could a release of PA or plasmin be demonstrated (see Fig. 1 for example). We also conclude from this experiment that the fibrinolysis observed with the 7-day rat paraflocculus is not due to damage (such as anoxia or nonspecific release of proteolytic enzymes). Extracellular release of PA and plasmin activity is even more obvious

if the conditioned medium from 7-day rat parafofloculus suspension culture is assayed using standard or heated fibrin plates. 6-Aminocaproic acid again inhibits the fibrinolytic activity and in these conditions neither PA nor plasmin activity could be demonstrated in medium conditioned by 1-month or adult rat parafofloculus.

We next investigated whether inhibition of PA or plasmin affected cell migration, using organotypic suspension cultures of the 7-day parafofloculus. In these cultures, the external granular layer (EGL) disappears within 2 days. Further experiments, including sequential autoradiography after intraperitoneal injection of ^3H -thymidine 60 min before dissecting the parafofloculus, showed that the disappearance of the EGL was due to inward migration of EGL cells such as occurs *in*

Fig. 1 Demonstration of plasminogen activator and plasmin release by 7-day-old rat parafofloculus. Fibrin plates were prepared as described in ref. 8;

200 mg of crude bovine fibrinogen (containing plasminogen) were dissolved in 100 ml of veronal buffer (pH 7.3) with 0.5 ml of 1% merthiolate solution added. To 3 ml of fresh fibrinogen solution (kept at 37°C), 0.7 ml of thrombin (100 NIH units per ml of Veronal buffer) was added and the mixture layered in the compartments of Petri dishes (Falcon 1004 Y). If the plates are kept at 37°C, both PA and plasmin activities can be detected. If the plates are heated to 80°C for 20 minutes, the plasminogen is inactivated and the tissue must generate not only endogenous PA but also endogenous plasmin in order to induce fibrinolysis⁹. The diminished lytic zone, as compared with standard plates, indicates that plasminogen in the tissue sample is the limiting factor. Parafofloculi were dissected from 7-day (a-c) or 1 month (d) rat cerebella. Meninges were carefully removed under the dissecting microscope. The whole (7-day-old) or part (1-month-old) parafofloculus was then explanted in 0.5 ml of N₁MEM medium serum-free culture medium¹⁵ on the fibrin plates which were incubated at 37°C for 24 h, and then photographed. Serum-free conditioned medium was obtained from suspension cultures of cerebellar parafofloculus in some experiments or of slices of the whole cerebellum in other experiments (see Fig. 2 legend for details on the tissue culture). The medium was collected, centrifuged (100,000g, 60 min) and filtered through a 0.22 µm Millipore filter. The conditioned medium was used immediately although activity was not modified by freezing at -20°C for 1 week. Forty µl of the conditioned medium were then seeded on either 37°C or 80°C fibrin plates, which were photographed after 24 h. a, 7-day-old rat parafofloculus, 37°C plate; b, 7-day-old rat parafofloculus, 37°C plate. 5×10^{-4} M 6-aminocaproic acid was added to the medium. At that concentration partial inhibition of the fibrinolysis is observed. Complete inhibition was obtained at 5×10^{-3} M; c, 7-day-old rat parafofloculus, 80°C plate; d, 1-month-old rat parafofloculus, 37°C plate; e, conditioned medium from 7-day-old rat parafofloculus suspension culture, 37°C plate; f, conditioned medium from 7-day-old rat parafofloculus suspension culture, 80°C plate. $\times 1.7$.

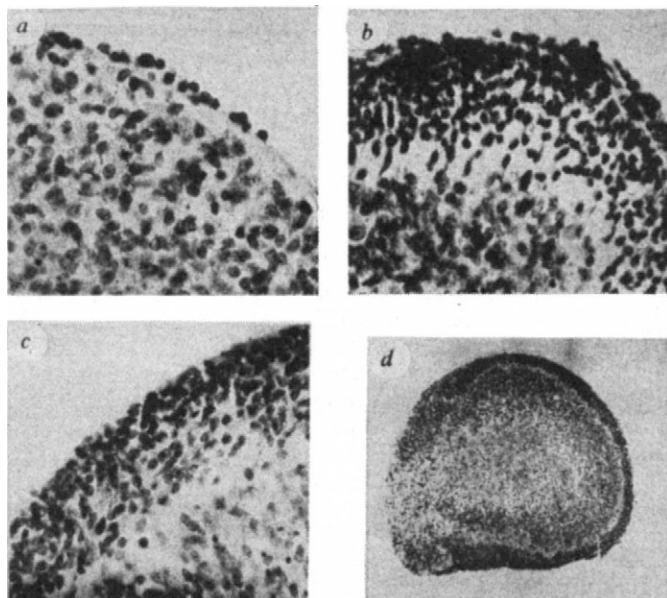
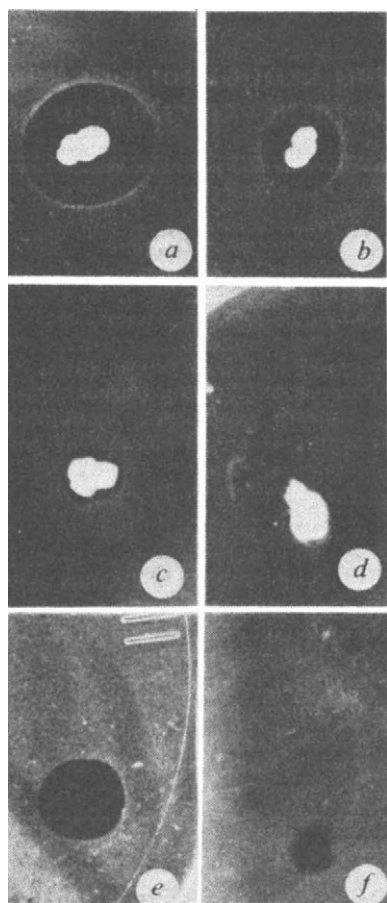


Fig. 2 Effects of inhibitors on neuronal migration. Cerebellar parafofloculi were dissected from 7-day-old rats and cultured in suspension (gyrotory shaker New Brunswick G2, 100 r.p.m.) for 48 h in high-glucose (600 mg%) minimal essential medium supplemented with either 10% heat-inactivated horse serum or with the N₁ component¹⁵ (the same results were obtained with both media). Different inhibitors of the PA-plasminogen-plasmin system were added (see text). a, Control culture showing the disappearance of the external granule cell layer after 2 days. b-d Illustrate the effect of some of the inhibitors used: b, 10^{-3} M EACA; c, 10^{-4} M D-Val-Phe-Lys-CH₂Cl; d, 10^{-4} M D-Phe-Pro-Arg-CH₂Cl. In the presence of inhibitor, the EGL does not disappear, showing that the inward neuronal migration is blocked. At least five samples were examined in each condition. Paraffin sectioning. Galloyanin stain. a-c, $\times 240$; d, $\times 42$.

vivo (in preparation). The persistence of the EGL beyond 2 days could thus be considered as a criterion for migration blockade. The following inhibitors were used: 6-aminocaproic acid, which binds to plasminogen and inhibits the transformation of plasminogen to plasmin; diisopropylfluorophosphate, a general inhibitor of serine proteases; trasylol and soybean trypsin inhibitors which are antiplasmin agents; and D-Phe-Pro-Arg-CH₂Cl and D-Val-Phe-Lys-CH₂Cl which are specific and potent inhibitors of tissue PA and plasmin respectively (ref. 10 and D. Collen, personal communication). All the inhibitors inhibited inward migration of EGL cells (Fig. 2), and, judging from phase-contrast examination of dissociated 7-day rat cerebellum cultures, at the concentration used they had no toxic effect. The data show that (1) not only PA but also plasmin activity are released into the extracellular space of the cerebellum during migration and (2) neuronal migration is inhibited by agents which act on PA and plasmin activity.

The present study is not concerned with cellular location of PA or plasminogen. Recent reports¹¹⁻¹⁴ suggest, however, that PA is released by neurones and, in the case of the 7-9-day postnatal mouse cerebellum, by granule neurones. The tissue localization of plasminogen is unknown, although there are several possibilities. For example, it could be present on the outer surface of the radial glia or released by the radial glia in the neighbouring extracellular space. Plasmin would be locally generated and the modification of cell surfaces or spaces on or near the glial process would generate an area permissive for inward migration of granule cells. This mechanism is consistent with the radial glia hypothesis. Alternatively, plasminogen could be distributed throughout the extracellular fluid and not restricted to the Golgi epithelial cell. Finally, it should be noted that the molecular nature of the substrate(s) for plasmin in the brain is at present unknown.

G.M. and I.S. acknowledge support from Fonds National de la Recherche Scientifique, Belgium.

Received 2 March; accepted 2 June 1982.

1. Swaiman, K. F. & Wright, F. S. in *Clinical Neurology* (eds Baker, A. B. & Baker, L. H.) 1-186 (Harper and Row, New York, 1977).
2. Sidman, R. L., Green, M. G. & Appel, S. M. *Catalog of the Neurological Mutants of the Mouse* (Harvard University Press, 1965).
3. Cajal, Ramon y S. *Histologie du Système Nerveux de l'Homme et des Vertébrés* Vols 1, 2 (Maloine, Paris, 1909-1911).
4. Miale, I. L. & Sidman, R. L. *Expl Neurol.* **40**, 277-296 (1966).
5. Rakic, P. *Trends Neurosci.* **4**, 240-244 (1981).
6. Reich, E. in *Biological Markers of Neoplasia: Basic and Applied Aspects* (ed. Rudson, R. W.) 491-500 (Elsevier, Amsterdam, 1978).
7. Quickley, J. P. *Cell* **17**, 131-141 (1979).
8. Astrup, T. & Mullertz, S. *Archs Biochem. Biophys.* **40**, 346-351 (1952).
9. Lassen, M. *Acta physiol. scand.* **27**, 371-376 (1952).
10. Collen, D., Lijnen, H. R., De Cock, F., Durieux, J. P. & Loffet, A. *Biochim. biophys. Acta* **165**, 158-166 (1980).
11. Soreq, H. & Miskin, R. *Brain Res.* **216**, 361-374 (1981).
12. Seeds, N. W., Haffke, S. C. & Krystocek, A. in *Tissue Culture in Neurobiology* (eds Giacobini, E., Vernadakis, A. & Shahar, A.) 145-154 (Raven, New York, 1980).
13. Krystocek, A. & Seeds, N. W. *Science* **213**, 1532-1534 (1981).
14. Krystocek, A. & Seeds, N. W. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7810-7814 (1981).
15. Bottenstein, J. E. & Sato, G. H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 514-517 (1979).

Transformation of fast fibres to slow prevented by lack of activity in developing lobster muscle

C. K. Govind* & Karla S. Kent†‡

* Scarborough College, University of Toronto, West Hill, Ontario, Canada M1C 1A4

† Boston University Marine Program, Marine Biological Laboratory, Woods Hole, Massachusetts 02543, USA

Adult mammalian twitch muscle fibres are extremely plastic, being able to transform from one type to another under the direction of their motoneurons¹⁻³. The most unequivocal type of fibre transformation is that of fast-twitch to slow-twitch type, either with cross-reinnervation of the fast muscle with a slow nerve⁴⁻⁶, or with chronic low frequency stimulation of the intact nerve to a fast muscle^{7,8}. Moreover, direct electrical stimulation of the muscle itself with a tonic pattern is just as effective in converting fast fibres to slow⁹. All these experiments suggest that activity of the muscle itself, imposed by whatever means, is crucial in the transformation of fast to slow fibres^{8,10}. A similar transformation during development may be responsible for producing the adult diversity of fibre types, as fibres begin by synthesizing myosin of an embryonic type which is replaced by a neonatal type and finally the adult type of fast¹¹ and slow¹², which occurs concurrently with innervation. We report here that during development of the paired claw closer muscle in lobsters, reducing or eliminating activity in one of the muscles by tenotomy or denervation prevents its fast fibres from becoming transformed into slow. Our results therefore suggest a regulatory role for activity in the differentiation of lobster slow fibres.

The adult lobster *Homarus americanus* has dimorphic claws consisting of a slender fast-acting cutter claw and a stout, slow-acting crusher claw^{13,14}. The claw closer muscles are similarly differentiated, the cutter having a majority (60-80%) of fast fibres and the crusher having all slow fibres^{15,16}. However, in the three larval¹⁷ and first juvenile¹⁸ stage the paired claws and closer muscles are symmetrical. Each muscle has ~30% fast fibres in a central band, with the remaining slow fibres distributed above and below the fast band¹⁹. During subsequent development, one of the muscles differentiates into a cutter with a majority of fast fibres and the other into a crusher with all slow fibres¹⁸. Claw and closer muscle type becomes determined some time during the 4th and early 5th stage and remains fixed thereafter²⁰. As claw laterality (whether the crusher appears on the right or left side) is not genetically prespecified, extrinsic factors may influence the determination of claw type

and muscle fibre type during the critical 4th and early 5th stage. Indeed, rearing juvenile lobsters without a substrate suppresses development of a crusher claw (and slow fibres) and instead produces paired cutter claws²¹. Because juvenile lobsters normally manipulate the substrate with their claws, it was proposed that the lack of exercise in the absence of a substrate suppressed the differentiation of a crusher claw. To test this hypothesis, we have examined the effects on claw dimorphism of rearing lobsters with or without substrates that they could manipulate and while restricting the use of one of the paired claws.

Larval lobsters were obtained from a hatchery and reared communally until the moult to the 4th stage, when they were placed individually in plastic trays²². Following experimental manipulation in the 4th and 5th stages, claw laterality was assessed in the 9th stage or beyond. At this stage the cutter claw is distinctly slender, covered with sensory bristles and has a major incisor-like tooth on its pollex. In contrast, the crusher is stout, has few sensory bristles and has a major molar-like tooth. The fibre composition of the claw closer muscles was based on histochemical demonstration of myofibrillar ATPase in frozen serial sections of the entire claw using the method of Padykula and Hermann²³, modified for lobster muscle²⁴. As the specific activity of myofibrillar ATPase is two to three times higher in fast crustacean muscle than in slow^{25,26}, the former fibres stained more intensely than the latter (Fig. 1). Usually,

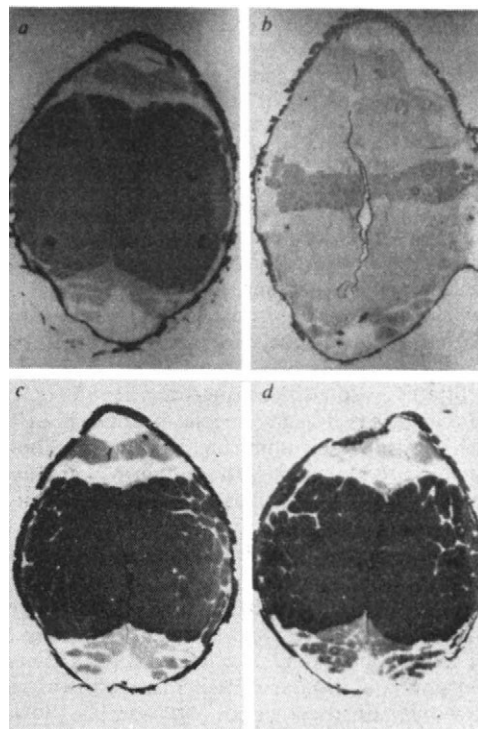


Fig. 1 Representative cross-sections midway through a serial series treated histochemically for myofibrillar ATPase in a pair of dimorphic cutter (a) and crusher (b) claws and a pair of cutter (c, d) claws from 9th stage juvenile lobsters reared with and without oyster chips, respectively. At this stage the closer muscle in the cutter claw (a) has almost completely differentiated, consisting mostly of intensely staining fast fibres with a small ventral band of less intensely staining slow fibres. Staining intensity of these ventral fibres is similar to that of the opener muscle which is known to be composed of all slow fibres. The closer muscle in the crusher claw (b) has a majority of fibres staining with the same intensity as the opener muscle thus denoting slow fibres except for a narrow central band of intensely staining fast fibres which are transformed into slow fibres in subsequent stages. In a juvenile lobster reared without a substrate, however, the closer muscles in the paired claws (c, d) are identical, having the fibre composition of a typical cutter muscle; a crusher muscle is completely suppressed. A detailed description of fibre types based on enzyme histochemistry in the lobster claw closer muscles has been published elsewhere¹⁹. $\times 18$.

‡ Present address: Department of Biological Sciences, Columbia University, New York, New York 10027, USA.

Table 1 Configurations of paired claws in juvenile lobsters reared with a variety of substrates and without any substrate

	Configurations of paired claws			Total no. reared	No. tested histochemically	P
	Right cutter and left crusher	Right crusher and left cutter	Double cutter			
Oyster chips	15	14	3	32	8	
Mud	7	12	3	22	5	NS
Gravel	13	7	2	22	5	NS
Plastic chips	13	9	1	23	8	NS
All substrate	48	42	9	99	26	
No substrate	1	0	33	34	9	<0.001

For each condition the lobsters were reared individually from the 4th stage until the 9th stage or beyond, when the configuration of the paired claws was assessed visually in all lobsters and histochemically in a few lobsters to determine the fibre composition of their closer muscles. The table lists the number of lobsters with each claw configuration, the total number reared and the number examined histochemically. With the oyster chips as the control condition, significance was assessed for each of the other substrate and no substrate conditions using the contingency χ^2 test. A cumulative tally of the different substrate-reared lobsters is given under the 'all substrate' condition. NS, not significant.

for each experimental condition 25% of the animals were tested histochemically (Tables 1, 2) and in each case the external claw type accurately reflected the fibre composition of its closer muscle. In other words, the cutter muscle consisted entirely of fast fibres except for some ventral slow fibres while the crusher possessed all slow fibres except for a central band of fast fibres (Fig. 1a, b). As the juvenile closer muscles were separated into these two distinct types, the external claw type served as a reliable indicator of the type of closer muscle within it.

We began by repeating the original experiment²¹ in which lack of a substrate suppressed development of a crusher claw and hence of slow fibres in 30% of the lobsters. In the original experiment juvenile lobsters were seen pinching the rubber stoppers over the drain hole with their claws. To prevent this, we covered the rubber stopper with a plastic tube and now found that 97% of the lobsters failed to develop a crusher claw (Fig. 1, Table 1). This was significantly different from the control condition of raising lobsters with oyster chips as substrate, which produced 91% with dimorphic claws and only 9% with both cutter claws. The two dimorphic configurations appeared in approximately equal proportions. Clearly, the lack of oyster chips as a substrate prevented transformation of fast fibres to slow and differentiation of a crusher muscle.

To test that it was not any special properties of the oyster chips which was the determining factor but rather their physical presence, we reared lobsters with mud, gravel and even plastic chips as substrate (Table 1). In each case, the majority (78.5%) of lobsters grew dimorphic claws; a minority failed to grow a crusher claw. For each of the different substrate conditions, the claw configurations were not significantly different from the control condition. Consequently, the presence of a substrate, of any kind, as long as it could be manipulated, is necessary for the differentiation of a crusher closer muscle. Furthermore, in the dimorphic condition, the crusher claw appears with almost equal probability on the right or left side, as shown when a tally is made of all substrate conditions (Table 1). This suggests that before the 4th and early 5th stage, each claw of a pair has

the potential for developing into a crusher but greater use of one claw in the critical period is the determining factor.

We therefore reared lobsters with oyster chips, but restricted the use of one of the claws to varying degrees. We arbitrarily chose to restrict use of the right claw while leaving the left one intact (Table 2). Keeping the claw immobilized in the closed position with a rubber band or cutting of most of its dactyl did not prevent differentiation of a crusher claw; the right, or restricted claw became a crusher as often as the left, or unrestricted claw. This suggests that use of the claw *per se* was not critical in the differentiation of the crusher claw but that possibly some other minimal condition such as muscle activity was responsible. With neither immobilization nor dactylotomy is activity of the closer muscle prevented.

Consequently, we attempted to reduce muscle activity using more drastic procedures such as deafferentation. This was done by painting either the entire claw or just the pollex and dactyl with a fast-drying lacquer. In the former case the dactyl usually remained in the closed position simply because the entire claw was painted whereas in the latter case it was free to open and close and was seen to do so. Both forms of deafferentation appeared at first glance to suppress differentiation of a crusher claw (Table 2) because the majority of the right claws were cutters. However, when tested statistically the number of right cutter claws in these two experimental groups was not significantly different from those in the control group. These forms of deafferentation with their consequent reduction in motor output were not sufficient to suppress development of a crusher claw and of slow fibres in the closer muscle.

We therefore eliminated activity of the closer muscle by cutting its tendon where it attaches to the movable dactyl, thereby disconnecting the muscle. Tenotomy of the right closer muscle effectively inhibited the treated claw from becoming a crusher compared with its untreated counterpart in the control group (Table 2). As tenotomy has several effects on the closer muscle, it is difficult to determine which particular effect suppressed transformation of fast to slow fibres. For instance,

Table 2 Form of the right claw—whether cutter or crusher, when untreated or treated in various ways

	Right claw		Total no. reared	No. treated histochemically	P
	Cutter	Crusher			
Control	18	14	32	8	
Immobilization	9	11	20	5	NS
Dactylotomy	8	7	15	5	NS
Paint claw	15	4	19	5	NS
Paint pollex and dactyl	15	3	18	5	NS
Closer tenotomy	34	3	37	9	<0.001
Opener tenotomy	19	3	22	5	<0.02
Denervation	27	5	32	10	<0.02

For each experimental group the right claw was manipulated in the 4th and again in the 5th stage and reared to the 9th stage, when its form was assessed visually in all lobsters and histochemically in a few lobsters for determining muscle fibre composition. Significance was determined between the control and each of the experimental conditions using the contingency χ^2 test.

tenotomy not only eliminates activity of the closer muscle but also eliminates its passive stretch and reduces the motor output to the claw muscles by disrupting the proprioceptive feedback.

On the other hand, tenotomy of the opener muscle which is an antagonist to the closer and is a small muscle occupying less than 10% of the claw muscle mass, also inhibited the treated claw from differentiating into a crusher (Table 2). Because this procedure allows the closer muscle to maintain its passive stretch, it rules out the possibility that loss of this stretch when the closer muscle was tenotomized suppressed the transformation of fast to slow fibres. Also, the relatively small size of the opener muscle would not seriously hinder contraction of the massive closer muscle when the opener is tenotomized. It is far more likely that tenotomy of the opener decreased motor output to the claw muscles because of reduction in proprioceptive feedback and sensory input from other sources. For example, some sensory nerve fibres from the dactyl run close to the insertion of the opener tendon and are presumably severed during tenotomy. Therefore, the failure to achieve some critical level of activity in fibres of the closer muscle prevents their transformation from fast to slow type.

As activity of the closer muscle is under the control of its nerve supply, we attempted denervation as a means of eliminating it without interfering with the muscle itself. This was done by destroying the anterior part of one side (right side) of the ganglion in intact animals. The cell bodies of the excitatory motoneurons to the claw closer and opener muscles are located in this region of the ganglion in adult lobsters²⁷. It was possible to destroy a relatively small anterior region of the ganglion with a finely sharpened pin. The efficacy of this surgery was confirmed by noting the loss of reflex activity in the claws and the destruction of the anterior part of the ganglion in post-mortem dissections of a few test lobsters. This form of denervation had a dramatic effect in that most lobsters failed to develop a crusher on the treated side (Table 2). Eliminating motor activity to the closer muscle prevents its fast fibres from transforming to slow.

Although denervation completely eliminates motor activity to the claw muscles, a reduction in activity such as with tenotomy of the opener muscle is just as effective in inhibiting formation of a crusher claw. Other, less acute treatments such as painting the claw, immobilization or dactylotomy, however, did not suppress development of a crusher claw, suggesting that a minimal level of muscle activity is required for the formation of a crusher claw. Our results therefore show that the reduction or elimination of muscle activity prevents the transformation of fast fibres to slow during development of the claw closer muscles in juvenile lobsters. Furthermore, the denervation experiments demonstrate the regulatory role of motoneurons in the differentiation of fibre types in lobster muscle. Such a crucial role for motoneurons in maintaining and determining fibre properties is well established for adult vertebrate muscle but has not previously been demonstrated in invertebrates and during development.

This work was inspired by the late Fred Lang. We thank H. L. Atwood for criticizing the manuscript, I. M. Campbell for help with statistical analysis, Joanne Pearce and Lena Hill for technical assistance and John Hughes of the Massachusetts State Lobster Hatchery on Martha's Vineyard for supplying us with larval lobsters. The work was supported by grants from the Natural Sciences and Engineering Research Council and the Muscular Dystrophy Association of Canada to C.K.G. and from the NIH (NINCDS) and the Graff Foundation to the late Fred Lang and A. G. Humes, Director, Boston University Marine Program.

Received 24 February; accepted 4 June 1982.

- Guth, L. *Physiol. Rev.* **48**, 645-687 (1968).
- Harris A. J. A. *Rev. Physiol.* **36**, 251-305 (1974).
- Gutmann, E. A. *Rev. Physiol.* **38**, 177-216 (1976).
- Buller, A. J. & Pope, R. *Phil. Trans. R. Soc.* **278**, 295-305 (1977).
- Close, R. J. *Physiol., Lond.* **204**, 331-346 (1969).
- Sreter, F. A., Luff, A. R. & Gergely, J. *J. gen. Physiol.* **66**, 811-821 (1975).

- Salmons, S. & Sreter, F. A. *Nature* **263**, 30-34 (1976).
- Salmons, S. & Henrikson, J. *Muscle Nerve* **4**, 94-105 (1981).
- Pette, D., Smith, M. F., Staudte, H. W. & Vrbova, G. *Pflügers Arch. ges. Physiol.* **338**, 257-272 (1973).
- Jolesz, F. & Sreter, F. A. *Rev. Physiol.* **43**, 531-552 (1981).
- Whalen, R. G. *et al. Nature* **292**, 805-809 (1981).
- Gauthier, G. F., Lowey, S. & Hobbs, A. *Nature* **274**, 25-29 (1978).
- Herrick, F. H. *Bull. U.S. Fish. Comm.* **15**, 1-252 (1895).
- Govind, C. K. & Lang, F. *J. exp. Zool.* **190**, 281-288 (1974).
- Jahromi, S. S. & Atwood, H. L. *J. exp. Zool.* **176**, 475-486 (1971).
- Lang, F., Costello, W. J. & Govind, C. K. *Biol. Bull.* **152**, 75-83 (1977).
- Lang, F., Govind, C. K. & She, J. *Biol. Bull.* **152**, 382-391 (1977).
- Govind, C. K. & Lang, F. *Biol. Bull.* **154**, 55-67 (1978).
- Ogonowski, M. M., Lang, F. & Govind, C. K. *J. exp. Zool.* **213**, 359-367 (1980).
- Emmel, V. E. *J. exp. Zool.* **5**, 471-484 (1908).
- Lang, F., Govind, C. K. & Costello, W. J. *Science* **201**, 1037-1039 (1978).
- Lang, F. *Aquaculture* **6**, 389-393 (1975).
- Padykula, H. A. & Hermann, E. J. *Histochem. Cytochem.* **4**, 161-169 (1955).
- Ogonowski, M. M. & Lang, F. *J. exp. Zool.* **207**, 143-151 (1979).
- Hajek, I., Chari, N., Bass, A. & Gutmann, E. *Physiol. Bohemoslov.* **22**, 603-612 (1973).
- Lehman, W. & Szent-Gyorgi, A. G. *J. gen. Physiol.* **66**, 1-30 (1975).
- Govind, C. K. & Lang, F. *J. exp. Biol.* **94**, 329-339 (1981).

Kainic acid stimulates excitatory amino acid neurotransmitter release at presynaptic receptors

John W. Ferkany, Robert Zaczek & Joseph T. Coyle*

Departments of Neuroscience, Pharmacology and Experimental Therapeutics, and Psychiatry and the Behavioral Sciences, Johns Hopkins University, School of Medicine, Baltimore, Maryland 21205, USA

Kainic acid (KA), a conformationally restricted analogue of glutamic acid, exhibits potent neuroexcitatory¹ and neurotoxic properties². The mechanism of the neurotoxicity of KA, however, seems to be complex and indirect because in many brain areas neuronal vulnerability requires the integrity of excitatory afferents³⁻⁶. Nevertheless, neurophysiological studies indicate that the neuroexcitatory effects of KA in the mammalian brain are direct⁷. We have now examined the effects of KA and other excitatory amino acids on the stimulation of cyclic GMP formation in brain slices incubated *in vitro* as a method for monitoring their depolarizing effects⁸. We show in the adult mouse cerebellum that the excitatory amino acid antagonist, D- α -amino adipate (DAA)⁹, blocks the stimulation of cyclic GMP produced by N-methyl-D,L-aspartate (NMDA) but potentiates the effects of KA. Whereas KA causes a significant release of both aspartic (Asp) and glutamic (Glu) acids by a calcium-dependent process, NMDA is without effect; furthermore, the effects of KA on Glu release are markedly reduced in cerebellum deficient in granule cells. KA also releases Glu and Asp from hippocampal and striatal slices, indicating that this response is not unique to the cerebellum. The results are consistent with evidence that KA has direct excitatory effects on neurones⁷ but suggest that its potent neurotoxic action involves the activation of presynaptic receptors on glutamatergic and aspartergic terminals, thereby releasing Asp and Glu.

Male albino mice weighing 25-35 g (Blue Spruce Farms) were used in all experiments. Some mice had been treated with methylazoxymethanol acetate (25 mg per kg) by intraperitoneal injection on the day of birth; this causes a selective deletion of granule cells¹⁰. Immediately after decapitation the brain was removed, and the cerebellum was rapidly isolated and sliced mechanically at 275 μ m at 2 °C before being placed in oxygenated Krebs-bicarbonate buffer. Histological analysis of the slices indicated preservation of the neuronal components in the conditions of slice preparation and incubation used. The cerebellar slices were preincubated for 60 min at 37 °C in Krebs-bicarbonate buffer constantly bubbled with O₂-CO₂ (95:5), with three changes of buffer. Portions of the tissue were

* To whom correspondence should be addressed at the Department of Neuroscience.

Table 1 Effects of excitatory amino acids on cerebellar cyclic GMP

Condition	Cyclic GMP formed (% of control)
L-glutamate (10 mM)	198 ± 15 (9)*
N-methyl-D, L-aspartate (0.5 mM)	159 ± 10 (9)*
N-methyl-D, L-aspartate (10 mM)	234 ± 22 (4)*
α-Kainic acid (0.1 mM)	156 ± 15 (4)*
α-Kainic acid (0.5 mM)	163 ± 9 (11)*
D-α-aminoadipate (0.25 mM)	116 ± 10 (22)
N-methyl-D, L-aspartate (0.5 mM) and D-α-aminoadipate (0.25 mM)	122 ± 14 (9)
α-Kainic acid (0.5 mM) and D-α-aminoadipate (0.25 mM)	251 ± 23 (11)*†

Cerebellar slices prepared from adult mice were incubated in conditions described in the text with the concentrations of amino acids indicated in parentheses and the amount of cyclic GMP formed was measured. The results are expressed in terms of per cent ($\pm$ s.e.m.) of the cyclic GMP levels in control slices incubated in parallel; basal cyclic GMP levels were 5.4 ± 0.5 pmol per mg protein ($n = 33$). Statistical comparisons were made on the basis of absolute levels of cyclic GMP by ANOVA.

* $P < 0.001$ versus control; † $P < 0.02$ versus kainate alone.

then transferred to vials with 2 ml of fresh buffer containing the compounds of interest, and incubation was continued for an additional 15 min. The reaction was terminated by brief centrifugation, the supernatant was removed, 1 ml of 0.05 M Tris-HCl buffer containing 0.5 mM EGTA was added, and the tissue was immediately boiled for 10 min at 100 °C and then dispersed by sonication. The tissue extract was assayed for cyclic GMP content by the method of Steiner *et al.*¹¹. In some experiments, slices were perfused with oxygenated buffer and the fractions were collected as previously described¹². Amino acids in the medium were measured by HPLC fluorescent detection by the pre-column derivatization method of Hill *et al.*¹³.

Incubation of the cerebellar slices with 10 mM Glu or 10 mM NMDLA caused a doubling of cyclic GMP levels (Table 1). Co-incubation of the slices with DAA (0.25 mM) blocked the cyclic GMP formation induced by NMDLA (0.5 mM). Dose-response curves with KA indicated that maximal stimulation of cyclic GMP formation occurred at 0.1 mM, but this concentration caused only a 60% increase in the cyclic GMP levels. However, the addition of DAA, which antagonizes the effects of the granule cell excitatory neurotransmitter¹⁴, markedly potentiated the effects of KA on cyclic GMP formation. This observation raised the question of whether release of endogenous excitatory amino acids in the presence of KA might be attenuating the cyclic GMP response by promoting the neurotoxic effects of KA.

The possible influence of KA on excitatory neurotransmitter release was evaluated by measuring the amount of endogenous amino acids released in the medium during drug exposure (Table 2). With KA (0.5 mM), the concentration of Asp in the medium increased to $233 \pm 33\%$ ($\pm$ s.e.m.; $P < 0.001$; $n = 6$) of control and that of Glu increased to $304 \pm 33\%$ ($P < 0.001$) of control whereas the concentrations of glutamine, alanine, tyrosine and taurine did not differ significantly from control. Notably, γ -aminobutyric acid (GABA) was not detected in the medium in the presence or absence of KA. The net release of Glu and Asp stimulated by KA was comparable with that found with a depolarizing concentration of potassium (40 mM). In contrast, neither NMDLA nor dihydro-KA (an analogue with low affinity for KA receptors^{1,15}) significantly altered the concentration of the Glu or glutamine in the medium. As dihydro-KA is a more potent inhibitor of the sodium-dependent, high-affinity uptake process for Glu^{5,16} than is KA, this result indicates that the KA-induced accumulation of Glu and Asp in the medium did not result simply from inhibition of the uptake process. Tetrodotoxin blocked the depolarization-induced release of Glu and Asp caused by the 40 mM KCl medium whereas tetrodotoxin did not antagonize the effects of KA.

Thus, the effect of KA did not depend on the generation of action potentials nor did it result simply from an increased extracellular concentration of potassium, a reported effect of KA¹⁷. Nevertheless, incubation of the slices in calcium-free medium containing EGTA (1 mM) markedly inhibited the amount of Glu and Asp released by KA although basal release was increased somewhat in this condition; thus, the KA-induced release probably involved stimulus-secretion coupling.

To establish that KA was acting on a neuronal pool of Glu, we examined the effect of KA on Glu release from cerebellar slices deficient in granule cells. Treatment of newborn mice with methylazoxymethanol acetate, a potent alkylating agent, causes a marked and selective depletion of the cerebellar granule cells which are undergoing mitosis at the time¹⁰. In slices prepared from the granule cell-deficient adult mice, KA (0.5 mM) did not cause a significant release of Glu ($P > 0.10$); rather, the net release of Glu in the presence of KA was only $38 \pm 14\%$ of that observed in unlesioned cerebella from littermate controls ($n = 9$). Thus, the granule cells, a reputed glutamatergic system^{18,19}, seem to be the primary source of Glu released by KA. However, KA did cause a significant release of Asp ($P < 0.01$; $n = 9$) in the granulo-prival slices which was 91% of the amount observed in controls; this is noteworthy because Asp seems to be a neurotransmitter of one of the climbing fibre systems²⁰, which are spared by the granule cell lesion²¹.

The presynaptic effects of KA on excitatory amino acid terminals were not limited to the cerebellum. In perfused slices prepared from cerebellum, hippocampus and striatum, 1 mM KA significantly increased the efflux of Asp and Glu (Table 3). Notably, the ratio of Asp to Glu differed among the three regions, with a greater increase in the Asp than Glu efflux stimulated by KA in the hippocampus. As observed in cerebellum, KA did not cause a release of GABA nor did it affect the efflux of glutamine or alanine. Perfusion with Ca^{2+} -free medium containing EGTA (1 mM) inhibited the KA efflux of Asp and Glu by $> 80\%$ ($n = 4$; $P < 0.01$) in the hippocampus.

Results of these experiments suggest a resolution to the conundrum that KA has direct excitatory effects⁷ although its neurotoxic effects seem to be primarily indirect and require the functional integrity of excitatory inputs³⁻⁶. The cerebellum offers a particularly favourable region for examining these interactions *in vitro* because the glutamatergic granule cell-parallel fibre system is intrinsic to the region. The present results

Table 2 Release of endogenous glutamate and aspartate from cerebellar slices

Condition	Net release (nmol per mg protein)	
	Glutamate	Aspartate
α-Kainic acid (0.5 mM)	2.04 ± 0.33*	0.61 ± 0.15*
Potassium (40 mM)	2.48 ± 0.02*	0.54 ± 0.14*
N-methyl-D, L-aspartate (0.5 mM)	1.02 ± 0.50	NM
α-Dihydrokainic acid (0.5 mM)	(-)0.03 ± 0.09	0.01 ± 0.14
Tetrodotoxin (5 μM)	0.06 ± 0.12	(-)0.19 ± 0.07
Potassium (40 μM) and tetrodotoxin (5 μM)	0.13 ± 0.14	0.01 ± 0.09
α-Kainic acid (0.5 mM) and tetrodotoxin (5 μM)	2.61 ± 0.70*	1.49 ± 0.28*
α-Kainic acid (0.5 mM) and no Ca^{2+} and EGTA (1 mM)	0.59 ± 0.32	0.36 ± 0.19

Cerebellar slices prepared from adult mice were incubated in the conditions described in the text with the concentrations of drugs indicated in parentheses. The results are expressed in terms of the net amount of amino acids released into the medium in the presence of drug compared with control slices incubated in parallel. Values are the mean $\pm$ s.e.m. of five or more preparations. In no condition did the levels of Gln differ significantly from control. NM, not measurable due to interference of N-methyl-D, L-aspartate in medium.

* $P < 0.01$ versus control by Student's *t*-test.

Table 3 Effects of kainic acid on the efflux of aspartate and glutamate from perfused slices of mouse brain regions

Region	Aspartate (% of baseline)	Glutamate
Cerebellum	440 ± 79*	963 ± 205*
Hippocampus	1,000 ± 280*	724 ± 57*
Corpus striatum	177 ± 44†	355 ± 66*

Immediately after decapitation, the relevant brain regions of the mice were isolated by dissection at 4 °C and placed in ice-cold oxygenated Krebs-bicarbonate buffer. The tissue blocks were then sectioned at 275 µm. After rinsing in Krebs buffer, the slices were transferred to syringe holders and supported on a bed of Sephadex G-10 (~15 mg wet weight tissue per chamber). The chambers were individually perfused at a rate of 0.5 ml min⁻¹ with oxygenated Krebs buffer maintained at 37 °C in a Haake constant temperature circulator. After 60 min of perfusion, spontaneous release of endogenous amino acids stabilized at a constant low rate. Three 3-min samples were collected for baseline values, then buffer containing 1 mM KA was substituted and three 3-min samples were collected for measurement of endogenous amino acid in the buffer by the HPLC method. The results are the mean ± s.e.m. of 4–8 separate preparations.

* $P < 0.01$ versus baseline; † $P < 0.05$ versus baseline.

demonstrate that the activation of guanylate cyclase in cerebellar slices by KA is significantly enhanced by DAA, which antagonizes the effects of NMDLA. The potentiating effect of DAA on KA is compatible with the interpretation that the stimulation of KA receptors combined with the release of endogenous excitatory amino acids has direct neurotoxic effects in the cerebella Purkinje cells, thereby attenuating the cyclic GMP response. This interpretation is consistent with the conclusions of Foster and Roberts²², who described cooperative interactions between exogenous excitatory amino acids and kainate on cyclic GMP stimulation in cerebellar slices prepared from 8-day-old rats. However, our study demonstrates, for the first time, the existence of presynaptic receptors for KA on the excitatory terminals.

In support of this hypothesis, Nicklas *et al.*²³ have previously demonstrated that KA alone profoundly depresses ATP levels in cerebellar slices, indicative of neurotoxicity²⁴. Furthermore, KA stimulates the release of both Glu and Asp in a calcium-dependent, tetrodotoxin-insensitive manner consistent with its action at a presynaptic receptor on parallel fibre and climbing fibre terminals. Ligand binding studies have demonstrated a partial, presynaptic localization of the specific binding sites for ³H-KA on the granule cells as these sites are reduced by 40% in the granule cell-deficient cerebellum¹⁰. Axonal receptors for KA on dorsal root fibres have also been described by neurophysiological methods²⁵. In conclusion, KA seems to act at two sites, one postsynaptic, which accounts for its direct excitatory effects, and the other presynaptic on glutamatergic and aspartergic terminals, which causes release of these excitatory amino acids and may enhance its neurotoxic effects.

This research was supported by USPHS grants NS-13584, MH-26654, DA-00266, USPHS postdoctoral fellowship NS-06798 to J.W.F. and RSDA II MH-00125 to J.T.C.

Received 2 November 1981; accepted 8 June 1982.

1. Biscoe, T. J., Evans, R. H., Headley, P. M., Martin, M. R. & Watkins, J. C. *Br. J. Pharmac.* **58**, 373–382 (1976).
2. Olney, J. W., Rhee, V. & Ho, O. L. *Brain Res.* **77**, 507–512 (1974).
3. Coyle, J. T. & Schwarcz, R. *Nature* **263**, 244–246 (1976).
4. McGeer, E. G., McGeer, P. L. & Singh, K. *Brain Res.* **139**, 381–383 (1978).
5. Biziere, K. & Coyle, J. T. *Neurosci. Lett.* **8**, 303–310 (1978).
6. Kohler, C., Schwarcz, R. & Fuze, K. *Neurosci. Lett.* **10**, 241–24 (1978).
7. McLennan, H. *Neurosci. Lett.* **18**, 313–316 (1980).
8. Foster, G. A. & Roberts, P. J. *Life Sci.* **27**, 215–221 (1980).
9. Watkins, J. C. & Evans, K. H. A. *Rev. Pharmac.* **21**, 165–203 (1981).
10. Slewin, J., Johnston, M. V., Biziere, K. & Coyle, J. T. *Dev. Neurosci.* **5**, 3–12 (1982).
11. Steiner, A. L., Ferrendelli, J. A. & Kipnis, R. M. *J. biol. Chem.* **247**, 1121–1124 (1972).
12. Nelson, M. F., Zaczek, R. & Coyle, J. T. *J. Pharmac. exp. Ther.* **214**, 694–702 (1980).
13. Hill, P. W., Walters, F. H., Wilson, T. D. & Stuart, J. D. *Analyt. Chem.* **51**, 1338–1341 (1979).
14. Stone, T. W. *Br. J. Pharmac.* **66**, 291–296 (1979).
15. London, E. D. & Coyle, J. T. *Molec. Pharmac.* **15**, 492–505 (1979).
16. Johnston, G. A. R., Kennedy, S. M. E. & Twitichin, B. *J. Neurochem.* **32**, 121–127 (1979).
17. Evans, R. H. *J. Physiol., Lond.* **298**, 25–35 (1980).

18. Young, A., Oster-Granite, M., Herndon, R. & Snyder, S. H. *Brain Res.* **73**, 1–13 (1974).
19. Sandoval, M. & Cotman, C. *Neurosciences* **3**, 199–206 (1978).
20. Wiklund, L., Toggenburger, G. & Cuenod, M. *Science* **216**, 78–80 (1982).
21. Nadi, N. S., Kanter, D., McBride, W. J. & Aprison, M. H. *J. Neurochem.* **28**, 661–662 (1977).
22. Foster, G. A. & Roberts, P. J. *Neurosci. Lett.* **23**, 67–70 (1981).
23. Nicklas, W., Krespan, B. & Berl, S. B. *Eur. J. Pharmac.* **62**, 209–213 (1980).
24. Retz, K. C. & Coyle, J. T. *J. Neurochem.* **38**, 196–203 (1982).
25. Davies, J., Evans, R. H., Francis, A. A. & Watkins, J. C. *J. physiol., Paris* **75**, 641–654 (1979).

An opioid benzodiazepine

D. Römer, H. H. Büscher, R. C. Hill, R. Maurer & T. J. Petcher

Sandoz Ltd, Pharmaceutical Division, Preclinical Research, CH-4002 Basle, Switzerland

H. Zeugner, W. Benson, E. Finner, W. Milkowski & P. W. Thies

Kali-Chemie Pharma Ltd, D-3000 Hannover, FRG

Although the benzodiazepines have been extensively investigated since their discovery about 20 years ago¹, only recently have antagonists of the classical actions of benzodiazepines been described². We have now discovered a class of benzodiazepines which have lost all affinity for benzodiazepine binding sites but which have moderate to high affinity for opiate receptors. Depending on the substituents introduced, it is possible to obtain compounds with selective actions on μ , δ and κ subpopulations of opiate receptors. The compounds are effective analgesics in various animal models of pain, exhibit an opioid activity spectrum in various test systems, and show none of the typical effects of minor tranquillizers. Here we report that 1-methyl-2(3-thienylcarbonyl)-aminomethyl-5-(2-fluorophenyl)-H-2,3-dihydro-1,4-benzodiazepine (KC5103; tifiuadom), acts selectively on opiate κ -receptors.

It has been believed previously that 1,4-benzodiazepines are a unique class of anxiolytic, anticonvulsant and sedative drugs belonging to the group of minor tranquillizers¹ and indeed the recently described antagonist Ro 15-1788 (ref. 2) is specific for these actions on the central nervous system (CNS). It is well established that most of these clinically active substances bind with high affinity to specific benzodiazepine binding sites in the brain and probably exert their specific effects by modulation of neuronal transmission of γ -aminobutyric acid (GABA).

We have studied benzodiazepines having an alkyl group at position 2, in particular the effect that various changes to this alkyl group³ have on the affinities of these compounds for CNS receptors⁴. We synthesized the 2-methoxymethyl derivative metaciazepam (patent DE-PS 250937; ref. 5), which is a clinically effective anxiolytic⁶ although having only low affinity for benzodiazepine receptors. As an extension of this work we developed 2-acylaminomethyl-1,4-benzodiazepines with a variety of substituents on the phenyl rings and in the side chain (patent DE-OS 2952279; ref. 7). Such compounds are effective in the hot-plate test in mice but this analgesia is not typical of the muscle-relaxant pseudoanalgesia shown by diazepam⁸, rather the compounds exhibit affinity for opiate receptors and have lost their affinity for benzodiazepine binding sites.

In comparative binding studies on rat brain homogenates, tifiuadom (Fig. 1) displaces ³H-naloxone from its binding sites

Fig. 1 Structural formula of tifiuadom.

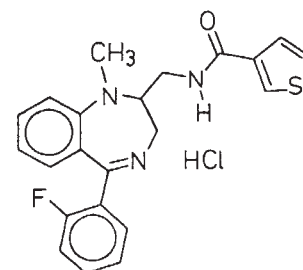


Table 1 Displacement of ^3H -(-)bremazocine from guinea pig brain homogenates

Substance	IC ₅₀ (nM)
Tifluadom	14 ± 6 (0.87)
Morphine	350 ± 170 (0.55)
Ketazocine	22 ± 12 (0.85)
Bremazocine	0.75 ± 0.5 (0.95)

Crude guinea pig membrane homogenates (minus cerebellum) were prepared essentially as described elsewhere¹⁶. ^3H -(-)bremazocine (0.5 nM) was incubated with membrane homogenates (17 mg per g wet weight per assay tube) for 40 min at ambient temperature in Tris-HCl buffer (50 mM, pH 7.4), filtered through Whatman GF/B glass fibre filters and washed twice with ice-cold assay buffer. All incubations were done in triplicate, and nonspecific binding was determined in the presence of 0.1 μM bremazocine. IC₅₀ values were calculated by appropriately weighted regression analysis (at least four concentrations per decade); Hill coefficients are given in parentheses.

with an IC₅₀ (50% inhibitory concentration) of 12 nM and has no effect on ^3H -flunitrazepam binding in concentrations up to 1 μM . In contrast, diazepam has no affinity for naloxone sites and displaces ^3H -flunitrazepam with an IC₅₀ of 15 nM. In the arthritis pain test in the rat⁹, a model of chronic pain processes in which classical benzodiazepines are known to give false positive reactions, tifluadom is effective with a 50% effective dose (ED₅₀) of 11 mg per kg orally (p.o.) and this action is partially reversed by pretreatment with 10 mg per kg naloxone subcutaneously (s.c.) but not by doses of up to 32 mg per kg intraperitoneally (i.p.) of the benzodiazepine antagonist Ro 15-1788. In the same test, diazepam is effective at a dose of 6 mg per kg p.o. and its action is readily abolished by Ro 15-1788 at 32 mg per kg i.p. but not by 10 mg per kg s.c. of naloxone. We conclude that tifluadom is an opiate analgesic with none of the actions typical of benzodiazepines.

Receptor binding studies with the universal opiate ligand ^3H -(-)bremazocine on guinea pig brain homogenates, a tissue preparation particularly rich in κ - and δ -opiate receptors¹⁰, suggested no particular preference for κ sites compared with μ and δ binding sites. Pharmacological evidence, however, suggested a preference for κ -receptors, therefore we performed comparative studies with the classical μ -agonist morphine, the prototypic κ -agonist ketazocine, tifluadom, and the highly potent κ -agonist bremazocine¹¹ (see Table 1). Table 2 summarizes the results obtained in electrically-stimulated isolated organs. The rabbit vas deferens (RVD) is unusually specific in its sensitivity to κ -agonists^{12,13}. Table 3 shows the effects of the four compounds in the mouse tail flick test¹⁴ and the reversal of these effects by the μ -antagonist naloxone and the putative κ -agonist Mr 2266 (ref. 15).

It is clear from these results that tifluadom is an effective opioid analgesic with a preference for opiate κ -receptors. In

Table 2 Effects on electrically-stimulated guinea pig myenteric plexus-longitudinal muscle preparation, and mouse and rabbit vas deferentia

Substance	IC ₅₀ (nM)			Potency ratio (normorphine = 1)	
	GPI	MVD	RVD	GPI	MVD
Tifluadom	1.7	2.6	4.9	234	43
Morphine	238	159	>10,000	0.9	0.8
Ketazocine	4.7	19.3	54	62	10.1
Bremazocine	0.25	0.5	6.4	1403	453

GPI, guinea pig myenteric plexus-longitudinal muscle preparation; MVD, mouse vas deferentia¹⁸; RVD, rabbit vas deferentia¹². The IC₅₀ is defined as the concentration required to produce 50% inhibition of the supramaximally electrically induced twitch response in each isolated tissue. Results are the means of several experiments: GPI $n = 6-12$, MVD $n = 6-10$, RVD $n = 5$. The absolute potency of normorphine varies strongly from tissue to tissue but is typically 150-400 nM (GPI) and 50-150 nM (MVD).

Table 3 Actions of test substances in the mouse tail flick test and their antagonism by naloxone and Mr 2266

Test substance	ED ₅₀ (mg per kg s.c., 30 min)	AD ₅₀ naloxone (mg per kg i.v.)	AD ₅₀ Mr 2266 (mg per kg i.v.)
Tifluadom	1.1 (0.5-2.4)	0.56 (0.31-1.02)	0.19 (0.08-0.45)
Morphine	3.0 (2.7-4.5)	0.005 (0.002-0.009)	0.08 (0.039-0.16)
Ketazocine	0.3* (0.1-0.5)	0.07 (0.03-0.1)	0.04 (0.02-0.08)
Bremazocine	0.7 (0.3-1.4)	~3.2	0.46 (0.24-0.87)

The ED₅₀ ($P = 0.05$) is defined as the dose that prolonged the reaction time in seconds by >75% in half of the mice. The AD₅₀ ($P = 0.05$) is defined as the dose (intravenous, i.v.) of the antagonist needed to reduce by 50% the analgesic effect of equi-active doses (ED₅₀) of the agonists. The antagonist was injected 15 min after the agonists (ketazocine, 7 min); readings were made 15 min later (ketazocine, 8 min): 5-10 animals were used per dose. Figures in parentheses indicate 95% confidence limits.

* 15 min.

keeping with these findings, the compound produces no respiratory depression in rats, is not self-applied by rhesus monkeys, and after programmed administration in the same species does not give rise to opiate withdrawal symptoms on cessation of treatment or after naloxone challenge. The existence of this new class of opioids suggests exciting possibilities for the study of opiate receptor subpopulations.

Received 7 June; accepted 29 June 1982.

1. Sternbach, L. H. *J. med. Chem.* **22**, 1-7 (1979); *Arzneimittel-Forsch.* **30**, 851-916 (1980).
2. Hunkler, W. *et al. Nature* **290**, 514-516 (1981).
3. Liepmann, H., Milkowski, W. & Zeugner, H. *Eur. J. med. Chem.* **11**, 501 (1976).
4. Li, P.-L. in *Proc. Workshop on Cell Surface Receptors—Basic and Clinical Aspects* Cambridge, 1982 (Horwood, London, in the press).
5. Milkowski, W. *et al. Chem. Abstr.* **86**, 89913W (1977).
6. Laakman G. & Hippus, H. *Adv. Biosci.* **31**, 125-134 (1981).
7. Zeugner, H. *et al. Chem. Abstr.* **95**, 169234b (1981).
8. Garattini, S., Mussini, E. & Randall, L. O. in *The Benzodiazepines* 36ff (Raven, New York, 1973).
9. Pircio, A. W., Fedele, C. T. & Bierwagen, M. E. *Eur. J. Pharmac.* **31**, 207-215 (1975).
10. Kosterlitz, H. W., Paterson, S. J. & Robson, L. E. *Br. J. Pharmac.* **73**, 939-949 (1981).
11. Römer, D. *et al. Life Sci.* **27**, 971-978 (1980).
12. Oka, T. *et al. Eur. J. Pharmac.* **73**, 235-236 (1981).
13. Römer, D., Hill, R. C. & Maurer, R. in *Proc. 3rd World Cong. on Pain*, Edinburgh, 1981 (Raven, New York, in the press).
14. D'Amour, D. E. & Smith, D. L. *J. Pharmac. exp. Theor.* **72**, 74-78 (1941).
15. Lord, J. A. H., Waterfield, A. A., Hughes, J. & Kosterlitz, H. W. *Nature* **267**, 495-499 (1977).
16. Pert, C. B. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2243-2247 (1973).
17. Kosterlitz, H. W. & Watt, A. J. *Br. J. Pharmac.* **33**, 266-276 (1968).
18. Henderson, G., Hughes, J. & Kosterlitz, H. W. *Br. J. Pharmac.* **46**, 764-766 (1972).

A novel intergenic regulatory element of prokaryotic operons

Christopher F. Higgins*, Giovanna Ferro-Luzzi Ames*, Wayne M. Barnes†, Jean Marie Clement‡ & Maurice Hofnung‡

* Department of Biochemistry, University of California, Berkeley, California 94720, USA

† Department of Biological Chemistry, Washington University School of Medicine, St Louis, Missouri 63110, USA

‡ Unité de Programmation Moléculaire et Toxicologie Génétique, CNRS LA271 INSERM U-163, Institut Pasteur, 75015 Paris, France

§ Present address: Molecular Genetics Laboratory, Department of Biochemistry, University of Dundee, Dundee DD1 4HN, UK

The nucleotide sequences of the intergenic regions from several multicistronic operons have recently been determined. In many cases such intergenic regions consist of only a few base pairs (bp); in certain extreme cases, where the initiation codon of one gene overlaps the termination codon of the previous gene, intergenic regions are essentially absent^{1,2}. However, in other cases intergenic regions can be up to several hundred base pairs long; it seems reasonable to suppose that, at least in these cases, such regions have a role in regulating expression of the operon. Indeed, a regulatory function for the large intergenic region of the *rpIL-rpoBC* operon has been demonstrated³. We have now identified a genetic element, common to intergenic regions from three independent

bacterial operons (the histidine transport and the histidine biosynthetic operons of *Salmonella typhimurium* and the *malK-lamB* operon of the *malB* region of *Escherichia coli*), which we believe to be important in regulating the differential expression of individual genes in these multicistronic operons. A similar element is also found immediately following the *trpR* gene of *E. coli* although in this case it is not known whether or not a distal gene exists. Each of these elements consists of a long dyad symmetry which, once transcribed, could form an exceptionally stable stem-loop structure with a stem of up to 35 bp. In each case smaller palindromic units are also found, some of which overlap the main symmetry. The stem-loop structures show remarkable homology (~90%) and probably have a common evolutionary origin, possibly an insertion element. It seems that the function of these structures is to effect a decrease in expression of distal genes in an operon.

The histidine transport operon of *S. typhimurium* consists of four cistrons, coordinately transcribed from a single promoter (see accompanying letter⁴). Between the first two genes of this operon, *hisJ* (encoding the periplasmic histidine-binding protein) and *hisQ* (encoding a membrane-bound protein), there is an intergenic region of 183 bp. An exceptionally long dyad symmetry (32 bases in each arm of the repeat) is evident within this region which, once transcribed, could form a very stable stem-loop structure ($\Delta G = -54.4 \text{ kcal mol}^{-1}$) (Fig. 1a). The *hisG-hisD* intergenic region of the histidine biosynthetic operon of *S. typhimurium* also contains a long dyad symmetry (Fig. 1c) which shows remarkable sequence homology with the dyad symmetry in the *hisJ-hisQ* intergenic region. A third example of this long intergenic dyad symmetry is found in the *malK-lamB* operon of *E. coli*⁶ (Fig. 1e) and again, it is almost identical in sequence to the two examples described above. This third example lies between *lamB* (encoding the λ

receptor) and *molA*, a gene whose function is unknown⁶. As no obvious promoter or terminator is found between *lamB* and *molA* it is reasonable to assume that these two genes form part of the same operon. A further example of this long dyad symmetry is found immediately following the *trpR* gene of *E. coli*^{7,8}. However, in this case it is not known whether the region is intergenic (that is, whether a distally located gene exists and is co-transcribed with *trpR*) or whether it marks the end of a transcription unit.

The remarkable homology between these dyad symmetries is shown in Fig. 2. For example, the long dyad symmetries in *hisJ-hisQ* and *hisG-hisD* show 90% homology (42/47 bases); a comparably high level of homology is evident whichever combination of regions is compared. Interestingly, the loops of these potential stem-loop structures show no significant homology.

Detailed analysis of sequences in and around these long dyad symmetries shows the presence of several copies of a smaller palindromic unit (Figs 2, 3). A computer search of sequence data banks did not reveal further copies of this unit at any other locus. These smaller palindromic units are not completely symmetrical (Fig. 3) and consecutive copies are always in opposite orientations. Thus, once transcribed, stem-loop structures could not only be formed within each unit (Fig. 1b, d, f), but could also form between units, giving rise to a variety of larger, secondary structures. Indeed, the intergenic elements (long dyad symmetries) described above are the result of just such base-pairing between two of these smaller palindromic units. Thus, stem-loop structures, some of which are mutually exclusive with the intergenic element, can be formed; such structures may or may not be functionally significant. Interestingly, the attenuator of the *rpLJL-rpoBC* operon³ also shows considerable homology with this smaller palindromic unit.

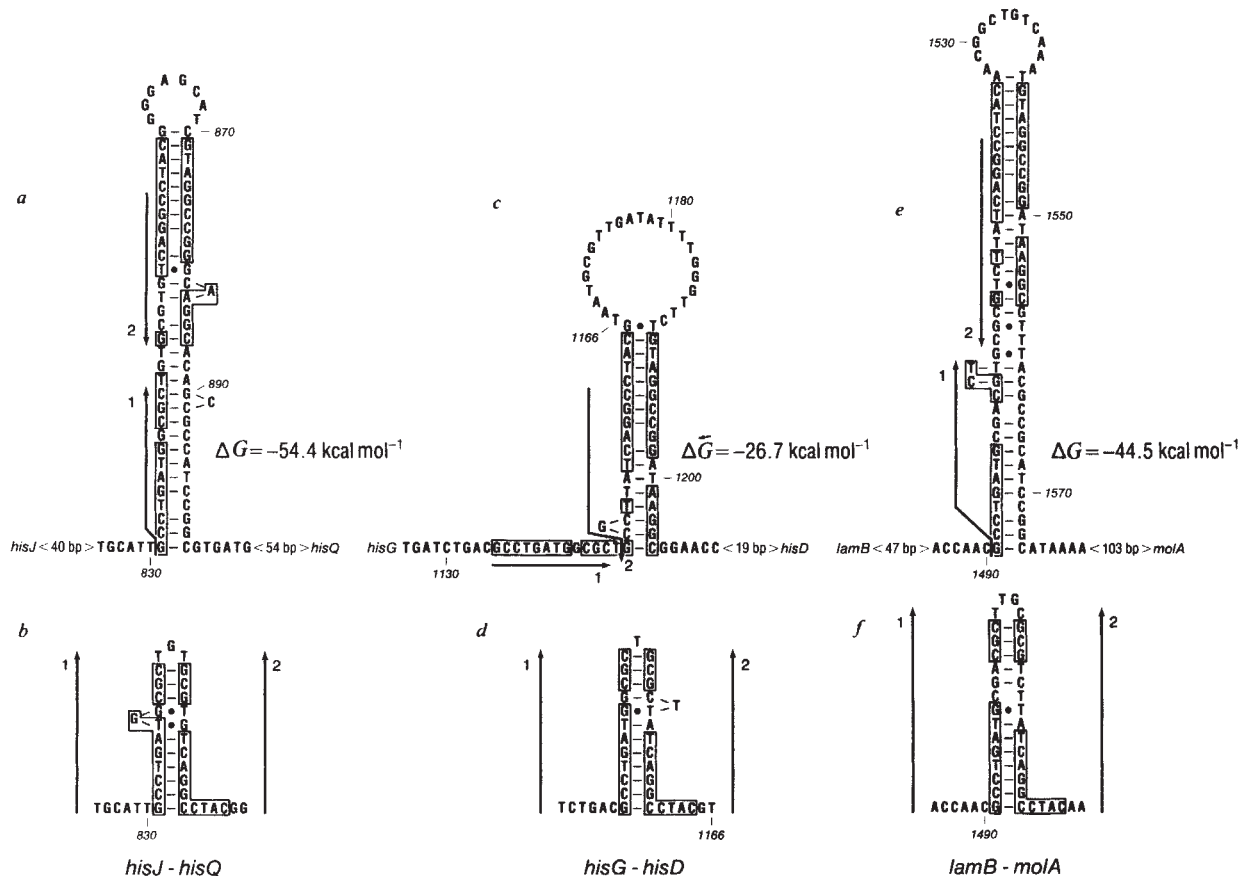


Fig. 1 Intercistronic regulatory elements. Sequences are found in: a, the histidine transport operon of *S. typhimurium* (*hisJ-hisQ*); c, the histidine biosynthetic operon of *S. typhimurium* (*hisG-hisD*); e, the *malK-lamB* operon of *E. coli* (*lamB-molA*). For each region alternative stem-loop structures are shown (b, d and f, respectively). The overlapping nature of these alternative structures is shown by the arrows (labelled 1 and 2) alongside the major structure which indicate the position of the alternative symmetry. Bases identical in all three sets of structures are boxed. Nucleotide sequences and base numberings are from ref. 4 (*hisJ-hisQ*), ref. 6 (*lamB-molA*) or unpublished data of W.B. (*hisG-hisD*).

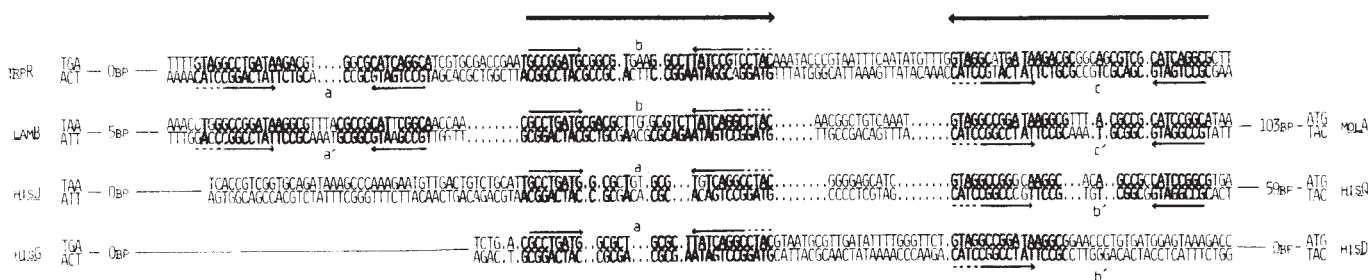


Fig. 2 Alignment of intercistronic regions. The *hisJ-hisG*, *hisG-hisD* and *lamB-molA* intercistronic regions together with sequences immediately following the *trpR* gene are aligned for homology. Shaded bases indicate significant homologies between the sequences, corresponding to the consensus sequence for the smaller palindromic unit (Fig. 3). A dot in the sequences indicates the absence of a base. Arrows above the sequences indicate the position of the long dyad symmetry (intercistronic element) discussed in the text. The shorter arrows above and below each sequence indicate the position of the smaller palindromic units. The homologous sequence 5' to each of these smaller palindromic units (see Fig. 3) is indicated by a series of dots following the arrows. Nucleotide sequences were obtained and are numbered as described in Fig. 1 legend or were taken from refs 7 and 8 (*trpR*).

What is the function of the intercistronic elements? Because of their location, we postulate that they are responsible for the lowered expression of distally located genes in a multicistronic operon. In the histidine transport operon, *hisJ* is expressed about 10 times more efficiently than *hisQ*⁴. Similarly, in the histidine biosynthetic operon, the dyad symmetry is found between two genes expressed to different extents: *hisG* is expressed at about three times the level of the distal *hisD* gene⁹. Unfortunately, in the *malK-lamB* operon the relative levels of expression of *lamB* and *molA* are unknown, although the *lamB* protein is known to be expressed at very high levels. It seems unlikely that regulation by these elements operates at a translational level: they are clearly separated from the ribosome-binding sites of the genes immediately downstream and bear no constant spatial relationship to the translation initiation codon. Indeed, evidence that regulation is transcriptional, rather than translational, has been obtained in studies of β -galactosidase production in strains carrying fusions of *lacZ* to *hisJ* and *hisQ* (M. Stern, C. F. H. and G. F.-L. A., in preparation). The elements might either cause premature termination of transcription, or they might serve as a processing site in the message after transcription. It is known that RNase III can cleave certain large stem-loop structures in mRNA¹⁰⁻¹² and it has been postulated that in doing so it regulates gene expression³. Perhaps, after RNase III cleavage, the distal portion of the message is rapidly degraded. It is also possible that these elements serve to protect long untranslated intercistronic regions from degradation, thus affecting the expression of distally located genes.

What is the origin of these elements? It seems unlikely that these homologous sequences have arisen independently to allow

specific interactions with a common nucleic acid sequence or protein: sufficient specificity could be imparted by very much shorter sequences. More probably, the extensive homology reflects a common evolutionary origin. It is interesting to speculate that these intergenic regulatory elements may have arisen from insertion sequence-like elements, that is, 'regulatory building blocks', becoming fixed at locations in which they provide a selective advantage. The lack of a short direct repeat on either side of the regulatory elements, typical of insertion sequences¹³, may be a result of mutational fixation of the element at these specific locations. Note, however, that these intergenic regulatory elements show no homology with known insertion elements of *E. coli*¹³.

Although we have identified three of these stem-loop structures in operons whose nucleotide sequences have been partially or completely determined, two lines of evidence indicate that very similar, if not identical, structures also occur at several other locations on the chromosome, presumably in other intercistronic regions where a step-down in expression of distal genes occurs. First, chromosomal rearrangements involving the histidine biosynthetic operon frequently have one end point within the *hisG-hisD* intergenic region¹⁴. It has previously been suggested that such *recA*-dependent rearrangements involve recombination with homologous sequences elsewhere on the chromosome¹⁴. Second, we have observed¹⁵ that a clone of the histidine transport operon of *S. typhimurium* hybridizes with a number of fragments from a total chromosomal digest, in addition to those fragments which cover the transport operon itself. This may be due to hybridization with homologous sequences dispersed throughout the chromosome. The precise number and location of these additional elements remain to be determined.

We thank Bruce Ames and D. Brutlag for many helpful discussions. This work was supported by NIH grants AM12121 and GM24956 to G.F.-L.A. and W.M.B., respectively, by a NATO/SRC fellowship to C.F.H., and by grants from the DDRST and the Fondation pour la Recherche Medicale to M.H.

Received 9 February; accepted 15 June 1982.

- Christie, G. E. & Platt, T. *J. molec. Biol.* **142**, 519-530 (1980).
- Selker, E. & Yanofsky, C. *J. molec. Biol.* **130**, 135-143 (1979).
- Barry, G., Squires, C. & Squires, C. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3331-3335 (1980).
- Higgins, C. F. *et al. Nature* **298**, 723-727 (1982).
- Borer, P. N., Dengler, B. & Tinoco, I. *J. molec. Biol.* **86**, 845-853 (1974).
- Clement, J. M. & Hofnung, M. *Cell* **27**, 507-514 (1981).
- Gunsalus, R. P. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7117-7121 (1980).
- Singleton, C. K., Roeder, W. D., Bogosian, G., Somerville, R. L. & Weith, H. L. *Nucleic Acids Res.* **8**, 1551-1560 (1980).
- Whitfield, H. J. *et al. J. molec. Biol.* **49**, 245-249 (1970).
- Gegenheimer, P., Watson, N. & Apirion, D. *J. biol. Chem.* **252**, 3064-3073 (1977).
- Robertson, H. D., Dickson, E. & Dunn, J. J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 822-826 (1977).
- Bram, R. J., Young, R. A. & Steitz, J. A. *Cell* **19**, 393-401 (1980).
- Calos, M. P. & Miller, J. H. *Cell* **20**, 579-595 (1980).
- Anderson, R. P. & Roth, J. R. *J. molec. Biol.* **119**, 147-166 (1978).
- Ardeshir, F., Higgins, C. F. & Ames, G. F.-L. *J. Bact.* **147**, 401-409 (1981).

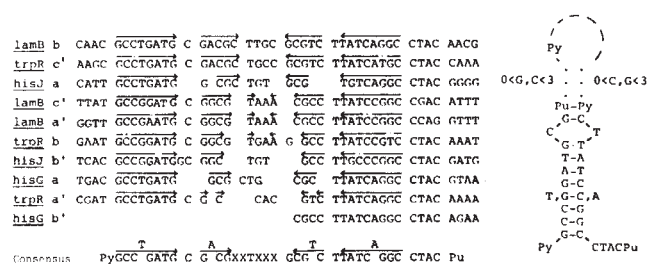


Fig. 3 Consensus sequence of palindromic units. The several copies of the smaller palindromic units identified in and around the intercistronic elements are aligned for maximum homology. Arrows above the sequence indicate regions of dyad symmetry. The location of these sequences with respect to each other is indicated in Fig. 2. A consensus sequence for these palindromic units is given: Py indicates a pyrimidine base, Pu a purine. Note that the units are not completely symmetrical; a conserved sequence, CTAC, is found 3' to each of the dyad symmetries, giving directionality to each unit. The potential stem-loop structure which might be formed by these units is shown schematically. Sequences were obtained as described in Figs 1 and 2.

Identification of a centriole-associated protein by antibodies present in normal rabbit sera

K. Turksen*, J. E. Aubin† & V. I. Kalnins*

* Division of Histology, Department of Anatomy and

† MRC Group in Periodontal Physiology, Medical Sciences Building, University of Toronto, Toronto, Ontario, Canada M5S 1A8

The centriolar region is a complex structural entity comprising not only centrioles, but also an abundance of associated electron-dense material identified at present only at the ultrastructural level^{1,2}. It is known to contain initiation sites necessary for the assembly of both cytoplasmic and spindle microtubules. Basal bodies, which are structurally related to centrioles, contain the template and serve as initiation sites for the assembly of other microtubule-containing structures, such as the axonemes of cilia and flagella. Immunoglobulins in the sera of some apparently normal rabbits have the ability to bind specifically to centrioles and basal bodies, and have allowed the distribution of these structures to be localized at the light microscope level by immunofluorescence³⁻⁵. The identity of the molecule(s) detected by these sera in the centriolar region is unknown, although it is not tubulin, the major identified protein of centrioles⁴. We describe here the identification of a protein recognized by the immunoglobulins in some normal rabbit sera, of approximate molecular weight (MW) 50,000 in both basal bodies isolated from *Tetrahymena pyriformis* and in extracts of chicken tracheal epithelial cells.

When a monolayer of mammalian cells is stained with one of these sera the centrioles are strongly labelled and their position in cells can be readily identified⁴. The staining of centrioles is not species-specific and can be seen in basal bodies and centrioles in a wide variety of species including protozoans such as *Tetrahymena* (Fig. 1a). When cells are treated with 0.5% Triton X-100 before fixation, the immunofluorescence staining of centrioles and basal bodies by these normal sera is abolished, suggesting that the centriolar antigen(s) is loosely bound to the centriole. Because it would be advantageous to use Triton X-100 to disrupt membranes and thus isolate the basal bodies from the pellicles of *Tetrahymena*, the basal body staining was stabilized by using EDTA (see Fig. 1b), a method previously used to stabilize the basal bodies in *Chlamydomonas*⁶.

A preparation enriched in basal bodies and retaining the centriolar antigen was made by breaking deciliated cells by vortex agitation in Triton X-100 containing 10 mM EDTA, pH 6.8, and separating the basal bodies from the cells and cellular debris by differential centrifugation (Fig. 2). Basal bodies were further purified by centrifugation on a discontinuous sucrose gradient. When basal bodies prepared in this way were stained for immunofluorescence with centriole-specific sera, the presence of the reacting antigen was clear. The protein components of the isolated basal bodies were then examined by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Numerous bands could be seen, including prominent ones at positions corresponding to molecular weights of 30–35,000, 45–47,000 and 55,000 and numerous less abundant ones at higher molecular weight (Fig. 2a). The band at about 55,000 apparently corresponded to tubulin, as it had the same electrophoretic mobility as the phosphocellulose-purified pig brain tubulin that was co-electrophoresed in a companion slot. This identity was confirmed by immunautoradiography of nitrocellulose⁷ sheets with antibody against pig brain tubulin.

To identify conclusively the antigen in the basal bodies which is responsible for the staining of the centrioles, we used the same electrophoretic blotting procedure. The basal body proteins were separated by SDS-PAGE (Fig. 2a) and electrophoretically transferred onto a nitrocellulose sheet (Fig. 2b).

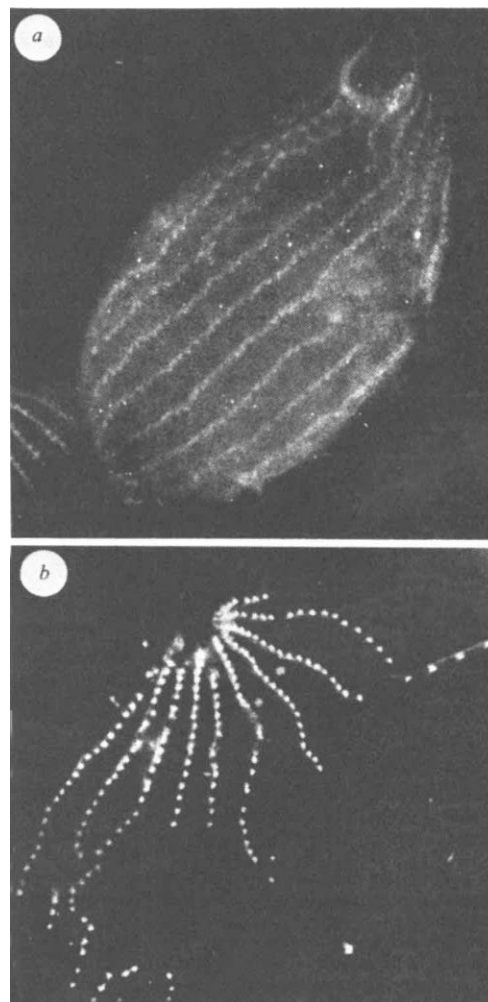


Fig. 1 Immunofluorescence micrograph showing typical basal body staining of isolated *Tetrahymena* pellicles by normal rabbit sera before (a) and after treatment with Triton X-100 + EDTA (b) stained with the same serum. Mass cultures of *T. pyriformis* strain GL cells were grown at 28 °C for 18–21 h in 6-l Erlenmeyer flasks, each with 1 l of culture medium containing 2% w/v protease peptone (Difco), 0.1% w/v liver extract L (US Biochemical). Cells were collected by centrifugation, and deciliated according to the procedure of Bird and Zimmerman¹⁰. Pellicles were isolated according to ref. 11. For immunofluorescence, purified pellicles were allowed to settle onto poly-L-lysine (Sigma, 0.1% w/v)-coated coverslips; they were fixed in methanol at –20 °C for 4 min, washed in phosphate-buffered saline (PBS) and then processed for indirect immunofluorescence (see below). For extraction, pellicles on poly-L-lysine-coated coverslips were fixed in freshly prepared 3% paraformaldehyde (Fisher) in PBS. The coverslips were then extracted in 1% Triton X-100 (Sigma)-containing 10 mM EDTA (Fisher) pH 6.8, for 5 min at room temperature and washed in PBS. They were then processed for indirect immunofluorescence. Standard indirect immunofluorescence: the coverslips after fixation were washed briefly in PBS, then incubated with normal rabbit serum (1:20 dilution in PBS) for 30 min at room temperature in a humidified chamber. The coverslips were then washed three times with PBS for 10 min and incubated for 30 min with fluorescein-labelled goat IgG against rabbit IgG (FITC-GAR, Hyland, diluted 1:5 in PBS) which was preabsorbed with *T. pyriformis* acetone powder. Finally, coverslips were washed three further times with PBS for 10 min. Coverslips were mounted in 50% glycerol in PBS, pH 7.8. The preparations were viewed with a Zeiss Photomicroscope-II equipped with epifluorescent illumination. Pictures were taken using Ilford FP-4 film. a, b, $\times 1,080$.

The sheets were then treated with several different centriole-positive normal rabbit sera and autoradiographed. With all the sera, one band could be readily detected in the autoradiographs (Fig. 2c) which, by comparison with marker molecules, corresponded to a molecular weight of 50,000 (Fig. 2). The band was not labelled with antisera to tubulin or with normal sera lacking specificity for centrioles (Fig. 2c), indicating that these sera did not contain any immunoglobulin against the 50,000-MW basal body protein. Treatment of the nitrocellulose sheets with ¹²⁵I-labelled protein A alone showed no binding between the

^{125}I -protein A and *Tetrahymena* basal body proteins (data not shown). Although the antisera to tubulin did stain the tubulin bands, this staining was clearly not due to the centriole-associated protein, as in companion gel slots the bands recognized by tubulin antisera and centriole sera were different. This was demonstrated further by labelling a nitrocellulose sheet containing basal body proteins simultaneously with both centriole-specific normal serum and tubulin-specific immune serum. This procedure clearly distinguished the tubulin band from the lower molecular weight (50,000) protein recognized by the centriole-specific serum (Fig. 2). Non-centriole-staining normal rabbit sera did not label any bands in these basal body extracts (not shown).

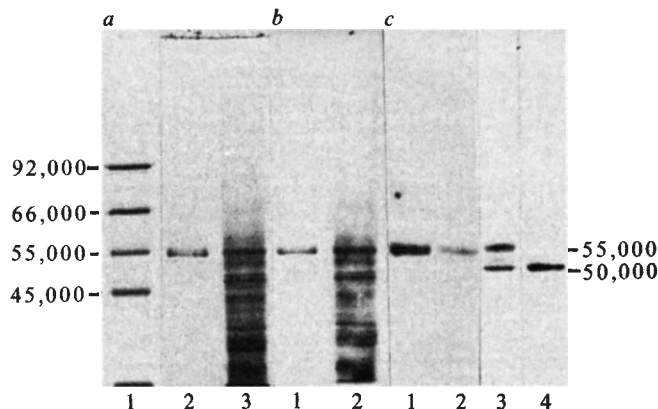


Fig. 2 Immunoelectrophoretic detection of the centriole-associated protein. *a*, SDS-PAGE stained with Coomassie blue. Lane 1, molecular weight standards phosphorylase *b* (92,000), bovine serum albumin (66,000), ovalbumin (45,000; Bio-Rad) and pig brain tubulin (57,000 and 55,000); 2, pig brain tubulin; 3, basal body-enriched fraction from *T. pyriformis*. *b*, Nitrocellulose sheet corresponding to gel in *a* after electrophoretic protein transfer and amido black staining. *c*, Immunautoradiogram from nitrocellulose sheets labelled with antibodies and ^{125}I -protein A. Lane 1, pig brain tubulin labelled with rabbit antiserum (1:750) to pig brain tubulin; 2, basal body-enriched fraction labelled with rabbit antiserum (1:500) to pig brain tubulin and with centriole-staining rabbit serum CT-2 (1:500); 3, basal body-enriched fraction labelled simultaneously with rabbit antiserum (1:500) to pig brain tubulin and with centriole-staining rabbit serum CT-2 (1:500); 4, basal body-enriched fraction labelled with centriole-staining rabbit serum HMW-3 (1:200). To isolate fractions enriched in basal bodies, deciliated cells were washed twice with fresh medium and broken in 1% Triton X-100 containing 10 mM EDTA, pH 6.8, by vortex agitation for 30 s. Unbroken cells and large debris were removed by centrifugation at 4,000g (Sorvall HB-4 rotor) at 4°C, for 10 min. The supernatant was further centrifuged at 8,000g for 10 min at 4°C. Finally, basal bodies were pelleted by centrifugation at 12,000g for 30 min at 4°C. This pellet is called the basal body-enriched fraction. In some experiments, the above pellet was run on a discontinuous sucrose gradient for further purification of basal bodies. For sucrose gradients, the basal body-enriched pellet was suspended in 1 ml of a 5% sucrose solution and was layered on a 0.7:1.2:1.7 M sucrose gradient (10 ml each) and centrifuged at 4,000g for 1.5 h using a Sorvall HB-4 swinging bucket rotor at 4°C. The band containing the basal bodies was at the 1.2:1.7 M interface. This band was removed carefully and diluted with distilled water. The basal bodies were pelleted by centrifugation at 12,000g for 30 min at 4°C. Basal body proteins were separated by SDS-PAGE using the buffer system of Laemmli¹². The samples were electrophoresed at a constant current of 30 mA, using a Bio-Rad slab gel apparatus model 220. A 3% stacking gel and 8% resolving gel were used. Basal body proteins were transferred to nitrocellulose sheets as described in ref. 7, and antigens were detected using ^{125}I -protein A. Briefly, the nitrocellulose sheets were saturated with 3% bovine serum albumin (BSA, Fraction V, Sigma) in 10 mM Tris-buffered saline (TBS), pH 7.4 by incubating for 1 h at 37°C. Sheets were then washed with TBS at room temperature and incubated for 24 h with rabbit sera in TBS buffer containing 3% BSA. The nitrocellulose sheets were washed for 1 h with TBS and incubated for 3 h with ^{125}I -protein A in TBS buffer containing 3% BSA. The nitrocellulose sheets were then washed in TBS containing 0.1% Tween 20 (Sigma) and dried in air. The blots were exposed to Kodak X-Omat XAR-5 film with an intensifying screen (Dupont) at -70°C. Protein A was iodinated at room temperature for 15 min in a reaction mixture (50 μl) containing 100 μg protein A (Pharmacia), 0.5 mCi carrier-free Na^{125}I (Amersham), 2.5 μl chloramine-T (Kodak), 5 μl dimethyl sulphoxide (Fisher). The reaction was stopped by addition of 5 μg of sodium metabisulphite (Sigma). The labelled protein and unreacted ^{125}I were separated on Sephadex G-25. The radioactively labelled protein A was rechromatographed on Sephadex G-25. Pig brain tubulin was purified as described¹³. The tubulin antiserum used in these studies was prepared against electrophoretically purified tubulin isolated from pig brain as previously described¹⁴.

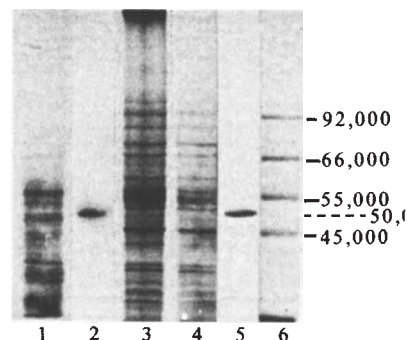


Fig. 3 Immunoelectrophoretic detection of the centriole-associated protein in ciliated chicken trachea cell extracts. Lane 1, nitrocellulose sheet with basal body-enriched fraction as in Fig. 2*b*, lane 2. Lane 2, immunautoradiogram corresponding to lane 1 but labelled with a centriole-staining serum CT-2 (1:200) and ^{125}I -protein A. Lane 3, SDS-PAGE of Triton X-100 extract of ciliated cells from chicken trachea. Lane 4, nitrocellulose sheet corresponding to gel in lane 3 after electrophoretic protein transfer and amido black staining. Lane 5, immunautoradiogram corresponding to lane 4 with nitrocellulose sheet labelled with centriole-staining rabbit serum CT-2 (1:200). Lane 6, molecular weight standards (see Fig. 2 legend). Ciliated cells were isolated from adult chicken trachea, obtained from a local abattoir, as described previously³. Briefly, the tracheae were rinsed with PBS and cut open at the midline. The tracheal epithelium was removed by gently scraping with a glass coverslip. Cells were collected by centrifugation at 4,000g (Sorvall HB-4 rotor) at 4°C for 10 min and extracted with Triton X-100 (Sigma) for 10 min. Cellular debris was removed by centrifugation at 12,000g for 10 min at 4°C. The supernatant was used in identification of the centriole-associated protein. The samples were electrophoresed and incubated as for Fig. 2.

Since the centriole-specific sera were known to label cells from diverse species of higher organisms, it was of interest to determine whether or not the same or a similar 50,000-MW protein could be detected in the centriolar structures in any of these organisms. Thus ciliated cells from chicken trachea were treated with Triton X-100, as the centriole staining of these cells was known to be lost by Triton X-100 extraction. When the detergent extracts of the ciliated cells were used for immunautoradiography, the centriole serum specifically labelled a band corresponding to MW 50,000 which comigrated with the antigenicity detected in basal body preparations (Fig. 3).

These results show that certain centriole-staining normal rabbit sera have IgG molecules directed against a protein with a mobility on SDS-PAGE corresponding to a MW of ~50,000. Similar IgGs are absent from sera which do not stain centrioles, or stain only the tubulin-containing component of centrioles.

The function—whether enzymatic or structural—of the 50,000-MW protein is unknown. However, it, or a very similar protein, seems to be present in centrioles and basal bodies of diverse cell types and organisms ranging from protozoa to mammals, suggesting that it has been conserved and may be functionally important. Recently, two other polypeptides (of MW 14,000 and 17,000) have been localized at the centriolar region using other preimmune or immune sera⁸. The use of these, and perhaps other, sera, as well as precise biochemical analyses (for example, ref. 9), may help to decipher important constituents of these complex structures which have not previously been defined. Finally, it would also be of interest to determine whether the ability to make these antibodies to centriolar proteins is genetically inherited or whether they are produced by the rabbits as a result of infection by basal body-containing microorganisms, or as a result of an immune response to centrioles, basal bodies or their constituents released by dying cells into the circulation.

We thank Dr A. Zimmerman for providing *T. pyriformis* strain GL. This work was supported by grant MT-3302 from the MRC of Canada to V.I.K.

Received 15 February; accepted 23 June 1982.

1. Roberts, K. & Hyams, J. S. (eds) *Microtubules* (Academic, New York, 1980).
2. Peterson, S. P. & Berns, M. W. *Int. Rev. Cytol.* **64**, 81–106 (1980).
3. Connolly, J. A., Kalnins, V. I., Cleveland, D. W. & Kirschner, M. W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2437–2440 (1977).
4. Connolly, J. A. & Kalnins, V. I. *J. Cell Biol.* **79**, 526–532 (1978).

5. Wang, E., Connolly, J. A., Kalnins, V. I. & Choppin, P. W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5719–5723 (1979).
6. Gould, R. R. *J. Cell Biol.* **65**, 65–74 (1975).
7. Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350–4354 (1979).
8. Lin, W., Fung, B., Shyamala, M. & Kasamatsu, H. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2373–2377 (1981).
9. Anderson, R. G. W. & Floyd, A. K. *Biochemistry* **19**, 5625–5631 (1980).

10. Bird, C. & Zimmerman, A. *Expl Cell Res.* **128**, 199–204 (1980).
11. Nozawa, Y. & Thompson, G. A. in *Biochemistry and Physiology of Protozoa* Vol. 2 (eds Levan Dowsk, M. & Hunter, S. H.) 275–338 (Academic, New York, 1979).
12. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
13. Weingarten, M. D., Suter, M. M., Littman, D. R. & Kirschner, M. W. *Biochemistry* **13**, 5529–5537 (1974).
14. Connolly, J. A., Kalnins, V. I., Cleveland, D. W. & Kirschner, M. W. *J. Cell Biol.* **76**, 781–786 (1978).

Mycoplasma pneumoniae adhesin localized to tip structure by monoclonal antibody

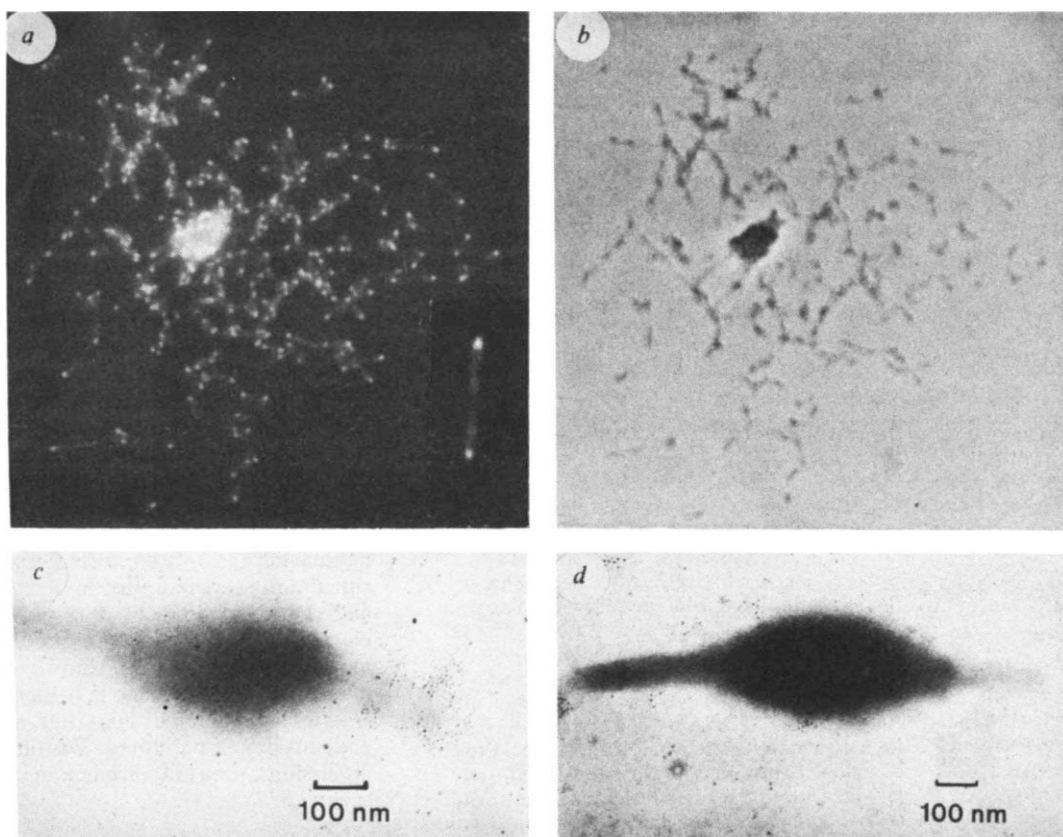
J. Feldner, U. Göbel & W. Bredt

Institute for General Hygiene and Bacteriology, Center for Hygiene, University of Freiburg, D-7800 Freiburg, FRG

The human respiratory pathogen *Mycoplasma pneumoniae* has several peculiar properties: despite its diminutive size¹ it possesses a specialized tip structure^{2–4}; it is motile and glides with considerable speed along inert surfaces^{5,6}, and it exhibits strong attachment to animal cells⁷ and to inert surfaces⁸. The latter property is a prerequisite for the survival of this obligate parasite in the host organism. Experimental evidence^{9–12} suggested that the binding site(s) mediating attachment of the pathogen is a protein, but so far this substance has not been characterized, nor has it been localized in the cell. By using a monoclonal antibody we have been able to localize the adhesin at the surface of the tip structure; it was identified by autoradiography as a protein of molecular weight (M_r) 160,000–190,000. We report that the monoclonal antibody inhibited erythrocyte adherence, attachment to glass and gliding motility of *M. pneumoniae*.

A hybridoma cell line was isolated according to the hybridization protocol described elsewhere¹³. Briefly, 6-week-old BALB/c mice were injected intraperitoneally with 0.5 ml of a *M. pneumoniae* (strain FH) cell suspension (1 mg ml^{-1}) in phosphate-buffered saline (PBS) pH 7.2. The injections were repeated four times at 2-week intervals. Mice were killed 3 days after the final injection, and their spleens removed. Lymphocytes were fused with Sp 2/0 myeloma cells¹⁴ and macrophages were used as feeder cells. The hybridoma clones were screened after 10 days for adherence-inhibiting properties by a haemadsorption method¹¹ and positive wells were cloned by limiting dilution. The supernatants were also tested by indirect immunofluorescence on *M. pneumoniae* cells grown on glass¹⁵. Of the 120 hybridomas examined, one clone was found to have adherence-inhibiting antibody, designated M43. On immunofluorescence, this antibody stained only one pole of the cell (Fig. 1a). The association of antibody with the tip structure of the *M. pneumoniae* cell was confirmed by electron microscopy using ferritin-labelled antibodies to mouse immunoglobulins (Fig. 1c). Controls treated with an unrelated mouse monoclonal antibody were negative (Fig. 1d). The immunofluorescence was reduced by pretreatment of the *M. pneumoniae* cells with trypsin ($10\text{--}20 \text{ } \mu\text{g ml}^{-1}$) at 37 °C for 30 min).

Fig. 1 Immunofluorescent and immunoferritin labelling of *M. pneumoniae* cells by monoclonal antibody M43. *a*, For indirect immunofluorescence the glass-attached mycoplasmas were fixed either in 0.1% glutaraldehyde or in 5% formaldehyde in PBS for 15 min at room temperature and rinsed twice with PBS. The cells were then treated with M43 hybridoma supernatant at 37 °C for 15 min, rinsed with PBS, and fluorescein isothiocyanate-conjugated goat anti-mouse anti-serum (1:40) was added as second antibody. Incubation at 37 °C for 15 min was followed by two rinses with PBS. Finally, the cells were covered with a buffered glycerol-phenylenediamine solution to reduce the fading of the fluorescence²¹. The preparations were examined in a Zeiss photomicroscope equipped with epifluorescence illumination. The photographs (3-min exposures) were taken using 400 ASA film (Ilford) and show a microcolony consisting of single cells ($\times 2,200$). The insert shows dividing cells ($\times 4,400$). *b*, Phase contrast of *a* without the insert. *c*, *M. pneumoniae* cells grown on parlodion-carbon-coated platinum grids were pre-fixed with 0.1% glutaraldehyde in PBS at room temperature for 15 min. After washing they were incubated with anti-adherence antibody (1:100) at room temperature for 15 min, then washed and incubated with ferritin-labelled goat anti-mouse immunoglobulin at room temperature for 15 min. The grids were then rinsed six times with PBS, fixed with 2.5% glutaraldehyde at room temperature for 15 min and examined in a Siemens Elmiskop 102 at 80 kV. Ferritin molecules are seen predominantly at the tip structure. *d*, As a specificity control, *M. pneumoniae* preparations were incubated with mouse monoclonal anti-human immunoglobulin instead of the M43 antibody. Further controls used bovine serum albumin, fetal calf serum or the ferritin-labelled antibody alone (not shown).



The antibody was identified as belonging to the IgM class by immunodiffusion and gel filtration on Sephacryl S-300. Pre-treatment of *M. pneumoniae* sheets with antibody M43 caused strong inhibition of adherence of sheep erythrocytes (Table 1). Similar results were obtained with human, rabbit and guinea pig erythrocytes. The adherence inhibition titre of hybridoma supernatant was 1:16; this increased to 1:256 when the antibody was precipitated with 40% ammonium sulphate and resuspended in PBS to 30 mg ml⁻¹. Similar titres (1:10 and 1:128, respectively) were found also for the inhibition of attachment to glass surfaces (Table 1).

Table 1 Monoclonal antibody M43 inhibits erythrocyte adherence and attachment to glass

	Adherence of sheep erythrocytes to mycoplasma sheet	Attachment of mycoplasmas to glass
Control	100	100
M43		
1: 32	3	1
1: 64	5	3
1:128	17	23
1:256	21	57
1:512	77	Not tested

Attachment of sheep erythrocytes was assessed by a haemadsorption technique¹¹ which allows photometric measurement of lysates of sheep erythrocytes attached to sheets of *M. pneumoniae*. Attachment of glucose-energized mycoplasmas to glass was quantitated in an experimental system^{12,20} which measures the settling of ³H-palmitic acid-labelled cells on to glass coverslips. After incubation and rinsing of the coverslips, the radioactivity attached to the glass was determined.

The monoclonal antibody caused no growth inhibition and was ineffective in the metabolic inhibition test¹⁶. Furthermore, no complement-mediated lysis of antibody-coated cells was observed. These negative results suggest a localization of the corresponding antigen at a non-vital site.

To identify the binding substance, the monoclonal antibody was labelled by adding ³⁵S-methionine to the hybridoma culture medium¹³. *M. pneumoniae* cells were solubilized by SDS and the proteins separated on a SDS-polyacrylamide gel¹⁷. The gel was then treated with the labelled antibody. Autoradiography revealed that the activity bound to a protein at a position

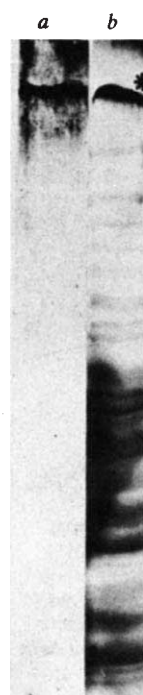


Fig. 2 Localization of the proteins stained by monoclonal antibody or by sheep erythrocytes after polyacrylamide gel electrophoresis¹⁷ of whole *M. pneumoniae* cells. After electrophoresis, the SDS was eluted from the gel by washing with PBS at room temperature for 45 min. *a*, The gel was then covered with ³⁵S-methionine-labelled antibody, incubated at 4 °C for 20 h, then washed with PBS. The fixed gel was impregnated with Enhance (NEN), then dried and exposed at -70 °C for 2 days to Kodak X-Omat X-ray film. *b*, Instead of antibody, sheep erythrocytes were used to overlay the gel (room temperature for 45 min). After several washes with PBS the gel was stained with Coomassie blue. The asterisk indicates the position of the erythrocyte-labelled protein.

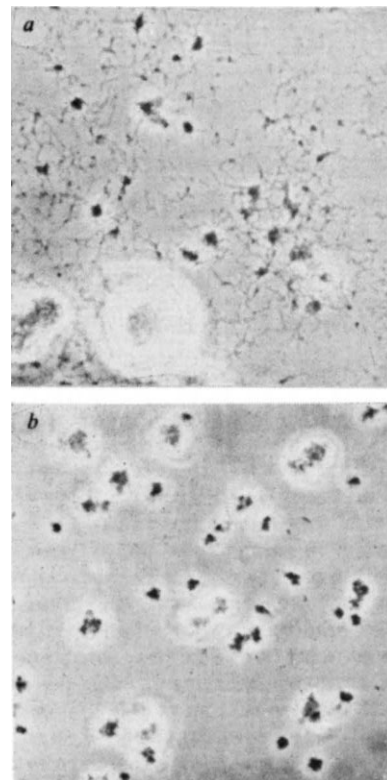


Fig. 3 Effect of antibody M43 on migration of *M. pneumoniae* cells. Mycoplasmas were grown on glass in coverslip chambers¹⁵ in the absence (*a*) or presence (*b*) of antibody M43 (1:3,200). In *b*, normal growth of microcolonies can be seen but no individual cells migrated out of the colonies.

corresponding to a molecular weight of 160–190K (Fig. 2*a*). The same protein band was reactive when a similar gel was overlaid with sheep erythrocytes (Fig. 2*b*), suggesting that this polypeptide was not only antigenically but also functionally identical to the binding site material.

The gliding movements of *M. pneumoniae* were completely abolished when antibody (up to 1:1,000) purified by ammonium sulphate precipitation was added to the medium. This inhibition occurred within 5 min; the shape of the cells was not affected nor did they lose contact with the glass surface. When the antibody was added to growing cultures, the single cells remained aggregated in clusters and did not migrate out of the microcolonies to spread over the coverslip (Fig. 3). The immobilization of the cells suggests that some mobility of the binding site within the membrane that is required for gliding, is inhibited by the bound antibody.

Attempts to isolate the adhesin protein of *M. pneumoniae* have so far had only limited success¹⁸. Using the specific monoclonal antibody described here it should be possible to purify and characterize the binding site protein, which may further elucidate the mechanisms of motility and adherence. Furthermore it could be used to study parasite–host interactions, and may have potential value in diagnosis and immunoprophylaxis.

After submission of this letter, an article describing a similar antibody¹⁹ was published. We thank A. Vogt for stimulating discussions, and D. Preusse and U. Burger for technical assistance.

Received 29 April; accepted 23 June 1982.

1. Razin, S. *Microbiol. Rev.* **42**, 414–470 (1978).
2. Biberfeld, G. & Biberfeld, P. *J. Bact.* **102**, 855–861 (1970).
3. Collier, A. M. in *Pathogenic Mycoplasmas* (eds Elliott, K. & Birch, K.) 307–327 (Associated Scientific, Amsterdam, 1972).
4. Göbel, U., Speth, V. & Bredt, W. *J. Cell Biol.* **91**, 537–543 (1981).
5. Radestock, U. & Bredt, W. *J. Bact.* **129**, 1495–1501 (1977).
6. Bredt, W. in *The Mycoplasmas* Vol. 1 (eds Barile, M. F. & Razin, S.) 141–155 (Academic, New York, 1979).
7. Razin, S., Kehane, I., Banai, M. & Bredt, W. *Ciba Fdn Symp.* **80**, 98–118 (1981).
8. Bredt, W., Feldner, J. & Kahane, I. *Ciba Fdn Symp.* **80**, 3–16 (1981).

9. Gorsky, F. & Bredt, W. *FEMS Microbiol. Lett.* **1**, 265–268 (1977).
10. Hu, P. C., Collier, A. M. & Baseman, J. B. *J. exp. Med.* **145**, 1328–1343 (1977).
11. Feldner, J., Bredt, W. & Kahane, I. *Infect. Immunity* **25**, 60–67 (1979).
12. Feldner, J., Bredt, W. & Razin, S. *Infect. Immunity* **31**, 107–113 (1981).
13. Fazekas de St Groth, S. & Scheidegger, D. *J. immun. Meth.* **35**, 1–21 (1980).
14. Shulman, M., Wilde, C. D. & Köhler, G. *Nature* **276**, 269–70 (1978).
15. Bredt, W. *Proc. Soc. exp. Biol. Med.* **128**, 338–340 (1968).
16. Senterfit, L. B. & Jensen, K. E. *Proc. Soc. exp. Biol. Med.* **122**, 786–790 (1966).
17. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
18. Banai, M., Razin, S., Bredt, W. & Kahane, I. *Infect. Immunity* **30**, 628–634 (1980).
19. Hu, P. C. *et al. Science* **216**, 313–315 (1982).
20. Feldner, J., Bredt, W. & Razin, S. *Infect. Immunity* **26**, 70–75 (1979).
21. Johnson, G. D. & de C. Nogueira Araujo, G. M. *J. immun. Meth.* **43**, 349–350 (1981).

A multicopy insertion sequence in the bovine genome with structural homology to the long terminal repeats of retroviruses

R. E. Streeck

Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France

Many of the transposable elements characterized so far in eukaryotes are a few thousand base pairs (bp) long and there are usually 20–50 copies per genome¹. In addition, families of short interspersed repeated sequences have been discovered, some of which may also be transposable². During analysis of the bovine satellite DNAs, I have found two types of sequence which have structural features in common with insertion elements³. They are only 0.6 and 1.2 kilobases (kb) long and are present in two different satellites at frequencies of 3.5×10^4 and 5×10^4 per genome, respectively. Sequence analysis of the 1.2-kb insertion and of satellite repeats carrying or lacking this sequence has now shown that a 6-bp duplication is generated at the site of insertion. I report here that the organization of this sequence is very similar to a retrovirus long terminal repeat (LTR). The 1.2-kb insertion and the related 0.6-kb sequence are highly homologous near their ends but unrelated in their internal sequences. These two elements may constitute precursors or descendants of an unknown retrovirus.

The bovine genome contains at least seven satellite DNAs, two of which have a density of 1.711 g cm^{-3} in CsCl (1.711A, 1.711B)⁴. Restriction and hybridization analysis of the 1.711B satellite (7.1% of the genome) has previously shown that it consists of two types of repeat unit: 1.4-kb units homologous to the 1.4-kb repeat units of another bovine satellite (1.715 g cm^{-3})^{5,6} and 2.6-kb units which correspond to 1.4-kb units carrying in addition a 1.2-kb insertion³. About 50% of the repeat units carry this insert, each at the same position. Thus, digestion of the satellite with *Sst*I, which cleaves once in the 1.4-kb sequence but not in the insert (Fig. 1a), yields mainly fragments of 1.4 and 2.6 kb. *Hind*III digestion, which cleaves once in the insert but not in the 1.4-kb sequence (Fig. 1a), gives mainly 2.6- and 4.0-fragments, while fragments ≥ 5.4 kb are less frequently found. This indicates that the 1.2-kb insertions occur most often in repeat units that are neighbouring or separated by one additional 1.4-kb unit, and with less probability in more distant ones.

I have determined the nucleotide sequence of the entire 2.6-kb repeat unit and have compared it to the 1.4-kb sequence derived from the 1.715 satellite DNA⁷. Uncloned restriction fragments were used for both sequences. They were therefore 'prototype sequences' corresponding to the most frequent nucleotides at each position of the repeat units. The sequence analysis confirmed that the 2.6-kb repeat unit consists of a 1.4-kb sequence highly homologous to the 1.4-kb repeat units, and an unrelated 1.2-kb sequence interrupting the 1.4-kb sequence. The homology between the 1.4-kb sequences in the repeat units carrying or lacking the 1.2-kb insertion is of the order of 95% (to be published). The nucleotide sequence of the 1.2-kb insertion (which is termed INS-1.711B) has several remarkable features:

(1) It is flanked by a 6-bp direct repeat (CCCTAC) which

occurs only once at the corresponding position in the 1.4-kb unit (Fig. 1b–d). Such a duplication is highly characteristic for the insertion of transposable elements¹.

(2) The insertion is bounded by a 3-bp terminal inverted repeat TGC...GCA (Fig. 2). If some mismatches are taken into account, the repeat could be extended to 13 bp.

(3) Following the terminal repeat, at the 5' end there is a repetitious sequence of ~100 bp consisting of several tri- to heptanucleotides repeated two to five times and interspersed with each other and some non-repeated DNA: for example, there are five copies of GCATATT. (These oligonucleotides are underlined in Fig. 2a.)

(4) An open reading frame of ~200 bp is found in the 5' half of the sequence. It is terminated by two successive TAA codons that are boxed in Fig. 2a.

(5) Several sequences thought to be important for the regulation of transcription are found in the 3' half of INS-1.711B (Fig. 2). These are a Goldberg-Hogness box, TATAAAAA, at positions 947–954 and a polyadenylation signal, AATAAA, at 966–971 which is followed 15 bp later by a polyadenylation site, Ca. The sequence GCAAG (900–904) precedes the Goldberg-Hogness box in a position where the related sequence CCAAT is frequently found. Finally, TTGGT (995–999) is found in a region where TTGT or related sequences are thought to have a role in the termination of RNA synthesis⁸ (Fig. 2).

The arrangement of these sequences in INS-1.711B is remarkably similar to that in a LTR of retroviruses. A comparison between INS-1.711B and the prototype of a retrovirus LTR is shown in Fig. 2b. A LTR contains sequences derived from both the 3' and 5' ends of the viral RNA (U3, U5) as well as one copy of the terminal redundancy (R), thus signals for the initiation and termination of transcription and for polyadenylation occur characteristically close to each other at fairly constant positions (Fig. 2b). In INS-1.711B the same or closely related sequences (see above) occur in these positions. Furthermore, the TG, CA-termini and the terminal inverted repeat present in all LTRs are also found in INS-1.711B. By analogy to retroviruses, the sequence 1–977 would then be U3-like, 978–987 would be R-like, and 988–1198 U5-like. It is not known, however, whether INS-1.711B is transcribed in bovine cells.

There are two structural features in INS-1.711B which have not been found in any of the sequenced LTRs: one is the repetitious region at the 5' end (Fig. 2a); the other is a 13 bp sequence (position 116–128), at the end of the repetitious

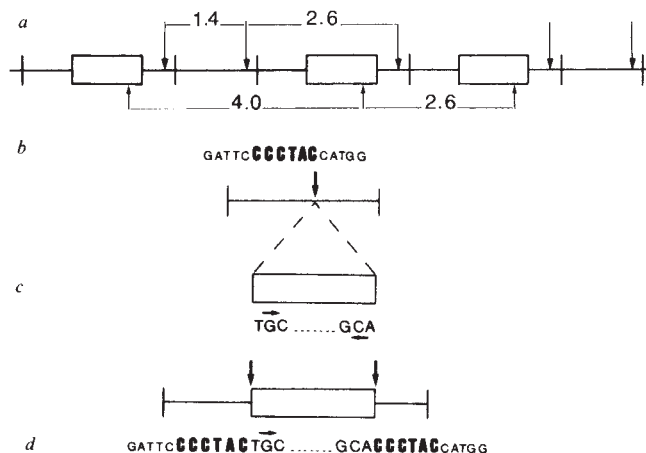


Fig. 1 Organization and partial sequences of repeat units in the 1.711B satellite DNA. *a*, Possible arrangement of 1.4-kb units lacking and 2.6-kb units carrying the 1.2-kb insertion. The sizes of fragments obtained by digestion with *Sst*I (↓) or *Hind*III (†) are given in kb. *b*, 1.4-kb repeat unit. The arrow indicates the site of insertion. The nucleotide sequence at this site is given⁷. *c*, 1.2-kb inserted sequence (INS-1.711B). The 3-bp terminal inverted repeat is indicated. *d*, 2.6-kb repeat unit carrying INS-1.711B. The sequences flanking the insertion and its 3-bp terminal inverted repeats are given, the duplicated hexanucleotide in bold letters.

region, that is related to the terminal repeats. In fact, this sequence forms a closer inverted repeat with the 3'-terminal sequence than does the 5'-terminal sequence (Fig. 2a).

No major differences were found when DNAs from thymus, liver and brain of different animals were digested by restriction nucleases and compared by Southern blot analysis⁹ using cloned INS-1.711B as a probe. Therefore, major rearrangements of INS-1.711B in the satellite DNA during ontogenesis can be excluded. Some background smears observed in the hybridization experiment suggest that INS-1.711B-like sequences are dispersed throughout the genome.

I have shown previously that a 0.6-kb sequence related to INS-1.711B occurs in the repeat unit of the bovine 1.711A satellite (termed INS-1.711A)³. It was therefore interesting to

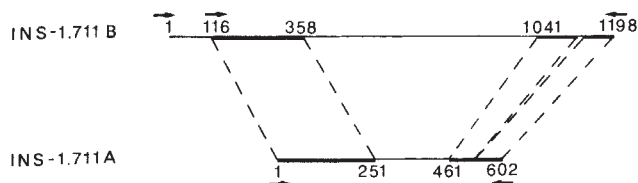


Fig. 3 Homologous regions of INS-1.711A and INS-1.711B. The nucleotide sequence of INS-1.711A is from ref. 3. INS-1.711A has been re-defined as the sequence bracketed by the inverted repeats, that is, the sequence 550–1,151 of the 1,413-bp repeat unit of bovine 1.711A satellite DNA³. The homologous sequences are indicated by bold lines, and their limits are represented by the corresponding nucleotide numbers. The arrows indicate the terminal inverted repeats and their internal repetition in INS-1.711B, respectively.

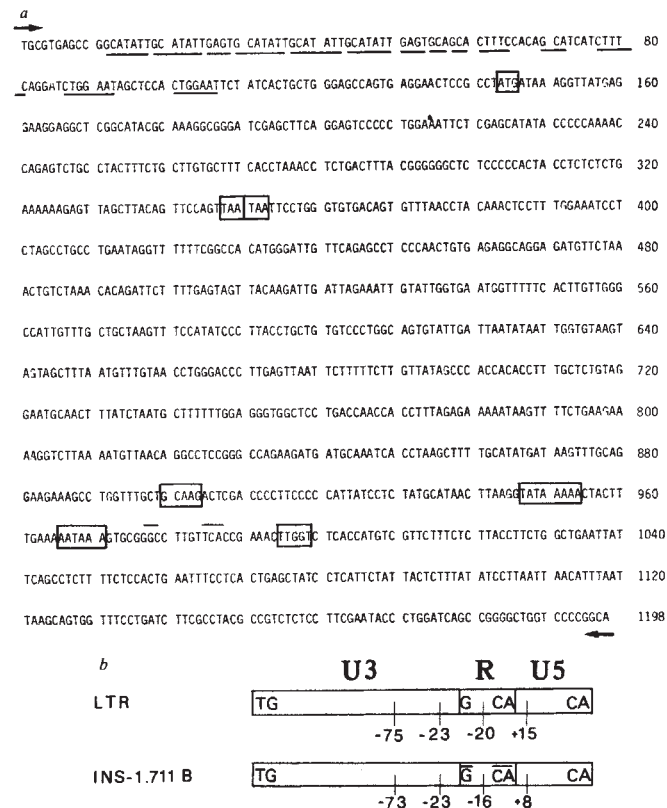


Fig. 2 Nucleotide sequence of INS-1.711B. Bovine 1.711B satellite DNA was purified from calf thymus DNA (Boehringer Mannheim) by isopycnic centrifugation⁴ analogous to that described for bovine 1.706 satellite DNA¹². Sequence analysis of restriction fragments from the purified satellite was carried out according to Maxam and Gilbert¹³, using the *AluI*, *DdeI*, *HindIII*, *HinfI*, *Sau3A* and *Sau96* sites in the insertion sequence for terminal labelling; 96% of the sequence was determined from at least two different overlapping restriction fragments, 40% being determined from both strands. *a*, Location of sequences mentioned in the text: the 3-bp terminal inverted repeats are indicated by arrows; the repeated oligonucleotide sequences near the 5' end are underlined; and the largest open reading frame is marked by the boxed ATG and the two successive TAA codons. The other boxes indicate regulatory signals and related sequences thought to be important in transcription and polyadenylation. They are, in succession, a CCAAT-like sequence, a Goldberg-Hogness box, a polyadenylation signal and a transcription termination signal. A possible adenylation site, CA, and a presumptive cap site, G, are overlined. *b*, Diagrammatic representation of a LTR⁸ and INS-1.711B. U3 and U5 correspond to sequences derived from the 3' and 5' ends, respectively, of the viral RNA; R indicates the terminal redundancy. The location of the regulatory signals mentioned above is marked. Their distances from the U3-R boundary (–75, –23) and R-U5 boundary (–20, +15)⁸ are given in bp. The locations of the homologous signals in INS-1.711B (boxed in *a*) are analogously indicated assuming that the sequence 978–987 (Fig. 2a) corresponds to R. The overlined G and CA in the R-like region correspond to those of *a*.

compare INS-711B with INS-1.711A, which also exhibits some features of an insertion element³. The homologous regions of the two sequences are the first 250 bp of INS-1.711A, which correspond to the sequence 116–358 in INS-1.711B, and the last 140 bp of both elements; their internal sequences are completely unrelated (Fig. 3). The homology is greatest near the ends, for example, two transitions being the only base changes in the last 35 bp at their 3' ends. The homology then gradually decreases to ~80% before ending abruptly. The 100-bp repetitive sequence found in INS-1.711B is absent from INS-1.711A. The region of homology near the 5' end which contains the largest open reading frame in INS-1.711B (Fig. 2a) is also conserved in INS-1.711A. The region of homology at the 3' end contains sequences characteristic of a polymerase II promoter in INS-1.711A³. In INS-1.711B this region is interrupted by a 19-bp sequence that is non-homologous. In summary, the homology found strongly suggests a common origin of the two sequences.

It is not known whether the high copy number of INS-1.711B is due to multiple insertion events or, more likely, to a passive amplification in the satellite DNA. In addition, it is not known whether INS-1.711B itself is transposable. Although the duplicated target sequence flanking the insertion is a strong indication of integration by transposition¹, it is also possible that INS-1.711B was a LTR of a larger movable element, for example, a retrovirus, which was first integrated and then excised by homologous recombination between the two LTRs (before being amplified in the satellite DNA). An analogous situation can be found for the solo δ -sequences of the Ty1 elements in yeast¹⁰. Bovine leukaemia virus, the only bovine retrovirus for which partial sequences are known (A. Burny, personal communication), is unrelated to INS-1.711B.

Alternatively, INS-1.711A and INS-1.711B may be different stages in the evolution of a cellular insertion sequence, which were conserved by (or possibly evolved through) their integration into repetitive sequences. According to Temin's hypothesis on the genesis of retroviruses¹¹, INS-1.711B might then be an immediate precursor of a retroviral LTR.

This work was carried out at the Institut für Physiologische Chemie, Physikalische Biochemie und Zellbiologie der Universität München, D-8000 München 2, FRG, and was supported by Deutsche Forschungsgemeinschaft. I thank A. Burny for the communication of unpublished results, D. Brutlag for help in using the computer facilities at Stanford University, and K. Beer for technical assistance.

Received 29 March; accepted 24 June 1982.

- Calos, M. P. & Miller, J. J. *Cell* **20**, 579–595 (1980).
- Singer, M. F. *Int. Rev. Cytol.* (in the press).
- Streeck, R. E. *Science* **213**, 443–445 (1981).
- Macaya, G., Cortadas, J. & Bernardi, G. *Eur. J. Biochem.* **84**, 179–188 (1978).
- Botchan, M. R. *Nature* **251**, 288–292 (1974).
- Philippsen, P., Streeck, R. E., Zachau, H. G. & Müller, W. *Eur. J. Biochem.* **57**, 55–68 (1975).
- Plucienniczak, A., Skowronski, J. & Jaworski, J. *J. molec. Biol.* **158**, 293–304 (1982).
- Temin, H. M. *Cell* **27**, 1–3 (1982).

9. Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
 10. Farabaugh, P. J. & Fink, G. R. *Nature* **286**, 352-356 (1980).
 11. Temin, H. M. *Cell* **21**, 599-600 (1980).
 12. Streeck, R. E. & Zachau, H. G. *Eur. J. Biochem.* **89**, 267-279 (1978).
 13. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).

DNA synthesis in purified DNA-membrane complexes extracted from a *Bacillus subtilis* pol A mutant

Patricia Benjamin, Paul Strumph, Mark Kenny & William Firshein

Department of Biology, Wesleyan University, Middletown, Connecticut 06457, USA

The cell membrane of prokaryotes has been implicated as the site of DNA replication in the replicon model first proposed by Jacob, Brenner and Cuzin¹. Ultrastructural, genetic and biochemical data have since been compiled suggesting an association between the cell membrane and replicating DNA (for review, see ref. 2). However, few studies have attempted to test this aspect of the replicon model directly by investigating the ability of DNA-membrane complexes to synthesize DNA *in vitro*. Such studies have been performed mainly with *Pneumococcus* and it has been shown that DNA-membrane complexes extracted by the M-band technique³ retain their capacity to synthesize DNA even after extensive purification⁴. However, mutants defective in various aspects of DNA replication are unavailable in *Pneumococcus*, and this is an important requirement for elucidating the physiological significance of the complex. Fortunately, they are available in *Escherichia coli* and *Bacillus subtilis*, and of the two organisms, much of the important work with the complex has been accomplished with *B. subtilis*⁵⁻¹². These studies have indicated that the origin and terminus of replication are permanently associated with the cell membrane. Nonetheless, there have been no previous studies of the synthetic capacity of *B. subtilis* DNA-membrane complexes. We now offer evidence that two DNA-membrane subcomplexes extracted and purified from a pol A mutant of *B. subtilis* are capable of synthesizing DNA *in vitro* and that these activities are inhibited by hydroxyphenyl azouracil, a specific inhibitor of DNA polymerase III activity in *B. subtilis*.

Many studies involving the DNA-membrane complex have been performed with only crude material, basically one step removed from whole cells. Because such complexes probably contain a variety of nonspecifically adsorbed components, it was considered essential to 'purify' the complex before analysing DNA synthesis. The purification procedure followed that developed for pneumococci⁴ because it fulfilled three important criteria. First, a large amount of excess macromolecules was dissociated during the purification procedure. Second, the complex retained its ability to synthesize DNA *in vitro* without added template or enzymes. Third, the specific activity of DNA polymerase(s) increased throughout the purification procedure (except in one step, see below). Table 1 shows the percentages of macromolecules remaining in the complex after the various purification steps. CsCl centrifugation separated membrane-bound DNA from the crude M-band³ into two peaks, one (corresponding to 94% of the total DNA in the crude complex) which sedimented at the buoyant density of free *B. subtilis* DNA, the other (corresponding to only 0.9% of the total DNA) floating at the top of the gradient as a pellicle together with 90% of the total lipid and 35% of the total protein.

These results indicate that DNA firmly bound to membrane represents a small fraction of the total DNA in the crude complex. One problem in this procedure is the possible contamination of the complex with 'free' protein which also floats at the top of the CsCl gradient. However, ~22% of the total protein of the crude complex was still associated with the

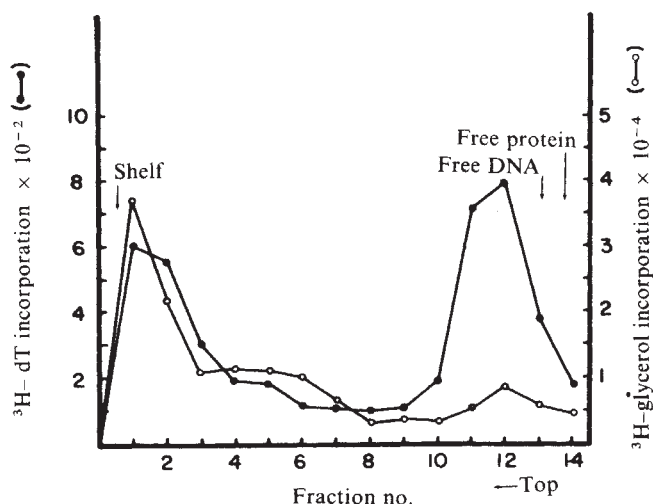


Fig. 1 Distribution of two DNA-phospholipid subcomplexes in a sucrose gradient. Methods are described in Table 1 legend. The experiment was repeated twice, once with a single label for DNA synthesis (³H-thymidine) and once with a single label for phospholipid synthesis (³H-glycerol).

subcomplexes after the final purification step described below. This suggests that at least this portion of the protein was firmly bound. The pellicle was dialysed to remove CsCl and centrifuged in a 5-20% (w/v) neutral sucrose gradient. Two peaks of DNA were observed (Fig. 1), one sedimenting near the bottom of the gradient and the other sedimenting near (but not at) the top. Both peaks contained varying amounts of co-sedimenting lipids (Fig. 1) and proteins (Table 1), suggesting that two DNA-membrane subcomplexes were resolved in the neutral sucrose gradient. Note that 'free' DNA and protein did

Table 1 Purification of the DNA-membrane complex

Treatment of complex	% DNA	Macromolecule protein	Remaining lipid
1 DNA-membrane dialysis	100	100	100
2 CsCl centrifugation, dialysis	0.86	35.6	90.0
3 Sucrose gradient centrifugation (total)	0.64	22.3	49.0
(a) More slowly sedimenting complex	0.28	18.9	9.1
(b) Faster-sedimenting complex	0.36	3.4	39.9

Cultures (500 ml) of *B. subtilis* strain 841 (thy, tryp, pol A) were labelled with 100 μ Ci of ³H-thymidine or 100 μ Ci of ³H-glycerol and grown to the mid-logarithmic phase in Penassay broth. Cells were washed twice with TMK buffer, pH 7.0 (0.01 M Tris, 0.01 magnesium acetate, 0.1 M KCl) and incubated in the same buffer containing 25% sucrose (w/v) and lysozyme (20 μ g ml⁻¹) at 4°C for 10 min. Protoplasts were collected, resuspended in TMK buffer containing 0.3% sarkosyl and 1 mM phenylmethylsulphonyl fluoride (PMSF), and incubated for 10 min at 4°C. The cell lysate was layered onto a biphasic 25/60% (w/v) sucrose gradient and centrifuged at 18,000 r.p.m. for 20 min at 4°C in an Sw 41 rotor and Beckman L2 65B ultracentrifuge. The M-band was collected from the 25/60% interphase and dialysed overnight in TK buffer (0.01 M Tris, pH 7.0; 0.1 M KCl) containing 1 mM PMSF at 4°C. Solid CsCl was added to the dialysate to obtain a mean density of 1.7001 at 30,000 r.p.m. for 48 h at 25°C in an Sw 50.1 rotor. The pellicle was removed from the top of each gradient, pooled and dialysed in TK buffer overnight at 4°C. The dialysate was layered onto 5-20% (w/v) neutral linear sucrose gradients containing an 80% sucrose shelf and centrifuged at 35,000 r.p.m. for 90 min at 4°C in an Sw 50.1 rotor. The gradients were fractionated by puncturing the bottom of the tube, each fraction precipitated with trichloroacetic acid (5%, final concentration) and assayed for radioactivity in a Mark III liquid scintillation counter. Samples were also taken from the pellicle fraction and other areas of the CsCl gradient for assay of radioactivity as well as from the M-band fraction. Protein content was determined by the Bio-Rad reagent.

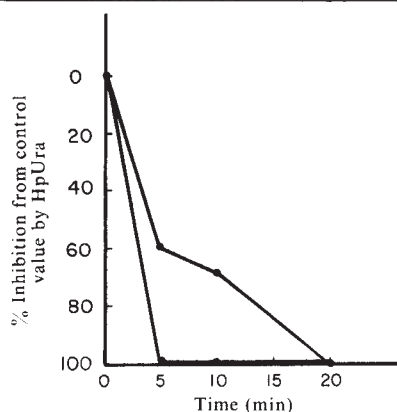


Fig. 2 Inhibitory effects of 6-(*p*-hydroxyphenyl) azouracil) on DNA synthesis by the two subcomplexes. Methods are described in Table 2 legend. The concentration of reduced Hpura (with sodium dithionite) was 400 μ M. The greater inhibitory activity (sharper slope) is observed in the fast-sedimenting subcomplex which is totally inhibited after 5 min, whereas the activity of the slowly sedimenting complex is totally inhibited only after 20 min. After 5 min the activity of the latter subcomplex is inhibited 60% from control values.

not sediment in the areas of the two subcomplexes but remained at or near the top of the gradient away from the slowly sedimenting subcomplex.

To determine whether the various complexes could synthesize DNA *in vitro* and whether purification of this activity was achieved, DNA polymerase assays were performed with the crude M-band, the CsCl pellicle fraction and the two subcomplexes detected in the sucrose gradient. Table 2 shows that both the total activity and the specific activity decreased after CsCl centrifugation and dialysis but increased significantly after sucrose gradient centrifugation. The reasons for this inhibition are unknown, but it has been observed consistently in our experiments. One possible explanation is that inhibiting factors are activated after this step but are removed after the subsequent sucrose gradient procedure. The more slowly sedimenting subcomplex showed a slight increase in specific activity while that of the faster-sedimenting complex was markedly

Table 2 DNA synthesis *in vitro* in various fractions during purification of the DNA-membrane complex

Fraction	Total activity (counts per min)	Specific activity (counts per min per mg protein)
1 DNA-membrane, dialysis	725,940	2,902
2 CsCl centrifugation, dialysis	52,425	526
3a Sucrose gradient centrifugation, more slowly sedimenting complex	119,040	3,737
3b Sucrose gradient centrifugation, faster-sedimenting complex	150,240	19,603

Cultures (41) of *B. subtilis* strain 841 were grown to the mid-logarithmic phase in Penassay broth in the same conditions as for Table 1 except that radioactively labelled precursors were omitted. The DNA-membrane complex was extracted and purified according to the conditions described in Table 1 legend. The fractions were assayed for DNA polymerase activity in a reaction mixture containing 1.7 mM triethanolamine phosphate buffer, pH 7.5; 0.5 mM EDTA; 0.025% Triton X-100; 2.0 mM β -mercaptoethanol; 0.01 M $MgCl_2$; dATP, dCTP, dGTP (25 μ M ml^{-1}); and 3H -dTTP (73 Ci $mmol^{-1}$), 1 μ Ci per assay tube. Each assay tube was incubated at 37 $^{\circ}C$ for 30 min and the reaction terminated with the addition of ice-cold 5% trichloroacetic acid containing 1% (w/v) sodium pyrophosphate. Acid-insoluble radioactivity was determined by liquid scintillation techniques.

Table 3 Transformation for genetic markers in the DNA-membrane complex

Transforming <i>B. subtilis</i> DNA	Transformants per ml cells		
	gua A1 ⁺ /leu 8 ⁺	pur A ⁺ /leu 8 ⁺	met B5 ⁺ /leu 8 ⁺
M-band-derived complex	20,000/445	413,000/104,000	40,000/5,500
CsCl-derived complex	600,000/10	90/40	38,240/610
Slowly sedimenting subcomplex	100,000/25	150/50	500/420
Fast-sedimenting subcomplex	120,000/20	60/30	660/80

Bulk *B. subtilis* DNA was extracted according to the method of Anagnos-topoulos and Spizizen¹⁷. Although not shown, the ratios of the various transformations for bulk DNA were similar to those of the M-band because the M-band contains ~80% of the total cellular DNA. The DNA membrane complexes were prepared as described in the legend (Table 1). These fractions were used at concentrations of 2.5 μ g ml^{-1} to transform recipient *B. subtilis* 1A177 (gua A1, ery-1), BUL717 (pur A, leu 8, met C), Mu8u5716 (pur A, met B5, leu 8) and a derivative of Mu8u5716 (gua A1, leu 8, pur A).

enhanced. It is interesting that the more active complex contained a higher concentration of lipids. Firshein and co-workers¹³ have shown that membrane glycolipids which are present only in Gram-positive bacteria stimulate DNA polymerase III activity in *Pneumococcus*. A similar phenomenon may be operating in *B. subtilis* if, in fact, the activity observed in both complexes is related to DNA polymerase III. Although it is probable that such is the case because of the pol A mutation (which virtually eliminates pol I activity), it was important to confirm this possibility by other means. Fortunately, 6-(*p*-hydroxyphenyl) azouracil (HpUra) specifically inhibits the activity of DNA pol III in *B. subtilis*¹⁴. Tests on the effects of this drug on the activity of the two subcomplexes show (Fig. 2) that with the addition of HpUra DNA synthesis is almost totally inhibited within 5–20 min, depending on the subcomplex, relative to controls.

Considerable evidence regarding the nature of membrane-associated DNA in *B. subtilis* has been provided by Sueoka and co-workers^{8,11,12} and Yoshikawa and co-workers^{7,9}. Their experiments have revealed several facts, which include (1) that the origin and terminus of replication are attached to the cell membrane throughout the cell cycle and (2) that the DNA-membrane complex can also be separated into two subcomplexes, both of which contain markers close to the origin. Although their subcomplexes were extracted and separated by different methods^{9,11,12} from those used in the present experiments, it was of interest to determine whether there was any similarity in the composition of the complexes, at least with respect to some of the genetic markers present. Accordingly, the ability of the DNA in each of the complexes to transform markers near the origin, terminus and middle of the *B. subtilis* genome was compared.

Table 3 shows that both the slow- and fast-sedimenting subcomplexes (as well as CsCl-derived complex) were enriched for a genetic marker (gua A1) which is located less than 1° from the origin of replication on the *B. subtilis* genetic map¹⁶. The commonly used origin marker (pur A), which has been mapped to ~10° from the replication origin¹⁶ and a genetic marker (met B5) which has been frequently used to detect enrichment for the terminus of replication, were not enriched in any of the DNA-membrane complexes isolated. In addition, no enrichment was detected for a marker (leu 8) located in the middle of the chromosome.

These results suggest that an unusually small area of the origin region may be represented in the subcomplexes. Although no enrichment for the terminus marker (met B5) was detected in the DNA-membrane complexes, it is possible that a smaller portion of the terminus region (including the actual terminus) is present in at least one of the subcomplexes. This suggestion is based on the possibility that the terminus marker (met B5) is positioned ~20° from the actual terminus of replication¹⁶, whereas only 0.64% of the DNA in the M-band complex

remains firmly bound in the subcomplexes (Table 1), a percentage corresponding to less than 2° of the *B. subtilis* chromosome. That this relatively small percentage does not represent a random area of the chromosome is indicated by the enrichment for the gua marker in most of the DNA-membrane complexes, but not for the pur A marker.

Finally, concerning the nature of the DNA synthesized, we can only surmise that since the assay solution does not contain components required for initiation of DNA replication (such as ribonucleoside triphosphates and high concentrations of ATP), an extension of strands from the replication fork is occurring with DNA pol III, rather than initiation of new strands. This possibility as well as attempts to determine whether initiation can be detected *in vitro* will be investigated. Nevertheless, this is the first report of a 'purified' DNA-membrane complex that can synthesize *B. subtilis* DNA *in vitro* using DNA pol III as the native enzyme.

This study was supported in part by NSF grant PCM 8110082 and by biomedical funds from Wesleyan University.

Received 29 March; accepted 14 June 1982.

- Jacob, F., Brenner, S. & Cuzin, F. *Cold Spring Harb. Symp. quant. Biol.* **28**, 329-348 (1963).
- Leibowitz, P. J. & Schaechter, M. *Int. Rev. Cytol.* **41**, 1-28 (1975).
- Tremblay, G. Y., Daniels, M. J. & Schaechter, M. *J. molec. Biol.* **40**, 65-76 (1969).
- Firshein, W. *J. molec. Biol.* **70**, 383-397 (1972).
- O'Sullivan, A. & Sueoka, N. *J. molec. Biol.* **69**, 237-248 (1972).
- Horowitz, S., Doyle, R. J., Young, F. E. & Streips, U. N. *J. Bact.* **138**, 915-922 (1979).
- Yamaguchi, Y. & Yoshikawa, H. *Nature new Biol.* **224**, 204-206 (1973).
- Beeson, J. & Sueoka, N. *J. Bact.* **139**, 911-916 (1979).
- Yamaguchi, Y. & Yoshikawa, H. *J. molec. Biol.* **110**, 219-253 (1977).
- Sueoka, N. & Quinn, W. G. *Cold Spring Harb. Symp. quant. Biol.* **33**, 695-705 (1968).
- Sueoka, N. & Hammers, J. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4787-4791 (1974).
- Imada, S., Carroll, L. E. & Sueoka, N. *ICN-UCLA Symp. molec. Cell Biol.* **3**, 187-200 (1975).
- Zerial, A., Gelman, I. & Firshein, W. *J. Bact.* **135**, 78-89 (1978).
- Mackenzie, J. M., Neville, M. M., Wright, G. E. & Brown, N. C. *Proc. natn. Acad. Sci. U.S.A.* **70**, 512-516 (1973).
- Kornberg, A. *DNA Replication*, 347-412 (Freeman, San Francisco, 1980).
- Henner, D. J. & Hoch, J. A. *Microbiol. Rev.* **44**, 57-92 (1980).
- Anagnostopoulos, C. & Spizizen, J. *J. Bact.* **81**, 741-746 (1961).

In vivo aminoacylation of brome mosaic and barley stripe mosaic virus RNAs

L. Sue Loesch-Fries* & Timothy C. Hall

Department of Horticulture, University of Wisconsin, Madison, Wisconsin 53706, USA

The genomic RNAs of several plant viruses are now known to be able to bind a specific amino acid at their 3' ends in a tRNA-like manner. Where the nucleotide sequence of the 3'-terminal region is known, it can be folded into an extensively base-paired secondary structure having the essential features of tRNAs^{1,2}. The high degree of nucleotide sequence conservation of the 3'-terminal regions of the four RNAs of the three known bromoviruses³, and the loss of infectivity of brome mosaic virus (BMV) RNA after chemical modification of its 3' end⁴, strongly suggest that the tRNA-like structure has a biological function in infection². To elucidate this it has been important to determine whether viral RNAs are aminoacylated within infected host cells. We studied BMV and barley stripe mosaic virus (BSMV), which belong to very different groups but which have genomic RNAs that are capable of both accepting tyrosine and infecting barley protoplasts. We report here that single-stranded RNAs of both viruses are aminoacylated *in vivo*, but that the encapsidated RNAs of BMV are not.

Both BMV and BSMV have segmented genomes. The molecular weights (MWs) of the genomic RNAs of BMV are 1.1×10^6 (RNA1), 0.99×10^6 (RNA2), and 0.75×10^6 (RNA3). RNA 4, a subgenomic RNA coding for coat protein, is encapsidated with RNA3 and has a MW of 0.28×10^6 (ref. 7). The RNA of the Type strain of BSMV (ATTC strain 43) used in

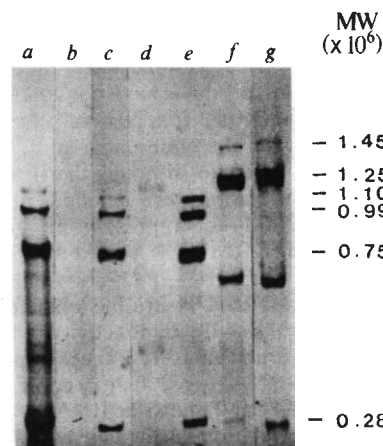


Fig. 1 Electrophoretic analysis of ³H-labelled RNAs from: a, BMV Tyr-RNA charged *in vitro*³; b, mock-inoculated protoplasts cultured in ³H-Tyr; c, BMV-infected protoplasts cultured in ³H-Tyr; d, mock-inoculated protoplasts cultured in ³H-uridine; e, BMV-infected protoplasts cultured in ³H-uridine; f, BSMV-infected protoplasts cultured in ³H-uridine; g, BSMV-infected protoplasts cultured in ³H-Tyr. BMV and BSMV were extracted from infected *Hordeum vulgare* L. cv. 'Dickson' or 'black hulls', respectively^{8,17}. Viral RNA was isolated by phenol extraction^{8,18}. Protoplasts from primary leaves of 'black hulls' barley (20 ml, 1×10^5 ml⁻¹) were mock-inoculated or inoculated with BMV or BSMV RNA at $5 \mu\text{g ml}^{-1}$ and cultured in the presence of ³H-Tyr ($30\text{--}60 \text{ Ci mmol}^{-1}$) at $30 \mu\text{Ci ml}^{-1}$ for 18–20 h, or ³H-uridine ($40\text{--}60 \text{ Ci mmol}^{-1}$) at $40 \mu\text{Ci ml}^{-1}$ and actinomycin D at $10 \mu\text{g ml}^{-1}$ for 24 h. RNA was phenol-extracted from the protoplasts using buffer containing 100 mM sodium acetate, 10 mM EDTA pH 4.5, 1% SDS and 1 mg ml⁻¹ bentonite for the samples labelled with ³H-Tyr, or buffer containing 10 mM glycine, 10 mM EDTA, 100 mM NaCl pH 9.5, 1% SDS and 1 mg ml⁻¹ bentonite for samples labelled with ³H-uridine. After ethanol precipitation, the RNA samples were suspended in water and stored at -80°C . Samples containing 2,000–3,000 c.p.m. were analysed by electrophoresis in $12.5 \text{ cm} \times 15.5 \text{ cm} \times 1.5 \text{ mm}$ slab polyacrylamide gels containing 2.4% acrylamide, 0.1% methylene bisacrylamide and 0.45% agarose, using 90 mM Tris-borate pH 8.3, 2.5 mM EDTA (TBE buffer). Electrophoresis was at 25 mA for 2 h at 4°C . The radioactive bands were detected by fluorography using EN³HANCE (NEN). Exposures of 30–50 days were required to detect radioactive bands. The MWs of the viral RNA bands are indicated.

this investigation can be resolved by gel electrophoresis into two RNAs of MW 1.4×10^6 (RNA1) and 1.25×10^6 (RNA2)^{8,9}. BMV and BSMV differ widely in their physical properties and transmissibility and thus have been assigned to separate virus groups; however, RNAs of both specifically accept tyrosine in an *in vitro* reaction catalysed by aminoacyl tRNA synthetase^{5,6}. *In vitro* studies have shown that aminoacylation of BMV RNA does not have a direct role in translation events^{10,11}; however, it may be necessary for RNA transcription or encapsidation^{12,13}.

Primary leaf protoplasts of barley were inoculated with BMV or BSMV RNA¹⁴; for mock-inoculation RNA was omitted. Protoplasts were cultured either in the presence of ³H-uridine and actinomycin D to label virus RNAs synthesized during infection, or in the presence of ³H-tyrosine (³H-Tyr) to label aminoacylated RNAs. Samples of RNA from inoculated protoplasts labelled with ³H-Tyr had approximately the same number of trichloroacetic acid (TCA)-precipitable counts as did a comparable sample from mock-inoculated protoplasts. However, striking differences in the patterns of aminoacylation became evident when the samples were analysed by gel electrophoresis (Fig. 1). Protoplasts inoculated with either BMV or BSMV RNA and cultured in ³H-Tyr (Fig. 1c,g) contained radioactive RNA species that co-migrated with the ³H-uridine-labelled viral species from infected protoplasts (Fig. 1e,f). The four radioactive bands of BMV RNA also co-migrated with BMV Tyr-RNA prepared by aminoacylation *in vitro* (Fig. 1a). RNA preparations from mock-inoculated (uninfected) protoplasts did not contain these labelled RNA bands. In addition to the genomic RNAs of BSMV, two smaller RNA bands were present in samples labelled with either ³H-Tyr or ³H-uridine. Low levels of these RNAs (MWs $\sim 0.65 \times 10^6$ and 0.30×10^6) are present in virion preparations; considerable amounts of them occur in infected protoplasts and they may represent subgenomic

*Present address: Agrigenetics Research Park, 5649 East Buckeye Road, Madison, Wisconsin 53716, USA.

messengers that are infrequently encapsidated¹⁵. They are also present in polysome preparations from infected barley; however, their translation products have not yet been identified (J. E. McFarland and A. O. Jackson, personal communication).

The ester bond between the amino acid and the viral RNA is readily hydrolysed. Therefore, although the viral RNAs from protoplasts cultured in ³H-Tyr were labelled, it was possible that the radioactivity did not represent an aminoacylated tyrosine. When BMV Tyr-RNA was heated in Tris-borate, pH 8.3 at 60 °C for 30 min, only 4% of the original radioactivity remained TCA-precipitable; if ³H-uridine labelled BMV RNA was similarly treated, 96% of the radioactivity remained TCA-precipitable. Electrophoretic analysis of these samples is shown in Fig. 2a-d. BMV Tyr-RNA (Fig. 2, lane c) contained an RNA that migrated as a broad band ahead of RNA4—this probably resulted from slight degradation during *in vitro* aminoacylation. This band is present in Fig. 2m, which also contains BMV Tyr-RNA. The RNA samples from BSMV- and BMV-infected protoplasts cultured in ³H-Tyr (Fig. 2e,f,i,j) behaved as expected for aminoacylated RNAs in that all radioactive viral bands lost most of their radioactivity during heating. Gels containing RNA samples from protoplasts cultured in ³H-Tyr subjected to electrophoresis for a shorter time (Fig. 2h-j) contained a rapidly migrating radioactive RNA band that appeared to be tyrosylated tRNA. This was confirmed by comparison of its electrophoretic migration with wheat germ tRNA aminoacylated *in vitro* (Fig. 2k,l).

RNA isolated from infected protoplasts includes viral molecules involved in translation and transcription as well as those encapsidated into virions. To determine whether encapsidated RNA was tyrosylated, virus was purified from infected protoplasts cultured in ³H-Tyr by a modified extraction procedure using differential centrifugation with unlabelled carrier virus¹⁴. RNA was isolated from the virus and analysed by polyacrylamide gel electrophoresis; RNA collected 18 h after inoculation contained only trace levels of radioactivity, but RNA from virus collected 24 h after inoculation contained

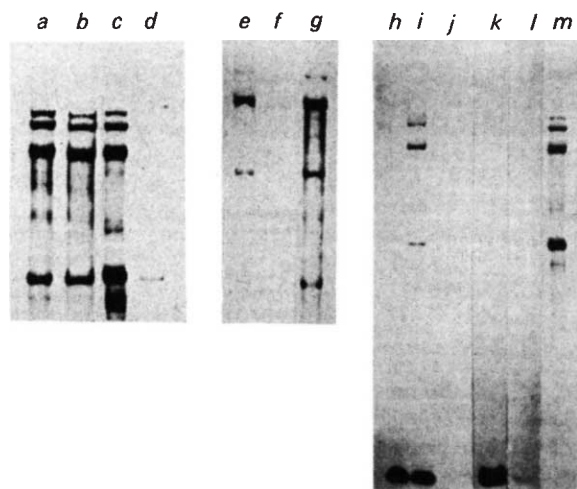


Fig. 2 Electrophoretic analysis of deacylated RNAs. Aliquots of RNA were prepared in either TBE buffer (see Fig. 1 legend) or in water as a control. The RNA samples suspended in TBE buffer were heated to 60 °C for 30 min and cooled on ice; samples suspended in water were brought to 50 °C for 5 min. All samples were made dense with urea. Electrophoresis was as described in Fig. 1 legend, except that RNAs in lanes h-m were subjected to electrophoresis for only 1.5 h. RNAs added to the gel lanes were: a and b, ³H-uridine-labelled BMV RNA from infected protoplasts; c and d, BMV Tyr-RNA aminoacylated *in vitro*; e and f, RNA from BSMV-infected protoplasts cultured in ³H-Tyr; g, RNA from BSMV-infected protoplasts cultured in ³H-uridine; h, RNA from mock-inoculated protoplasts cultured in ³H-Tyr; i and j, RNA from BMV-infected protoplasts cultured in ³H-Tyr; k and l, wheat germ Tyr-tRNA aminoacylated *in vitro*; m, BMV Tyr-RNA aminoacylated *in vitro*. RNAs in lanes b, d, f, j and l were heated in TBE buffer; the other lanes contain control RNA.

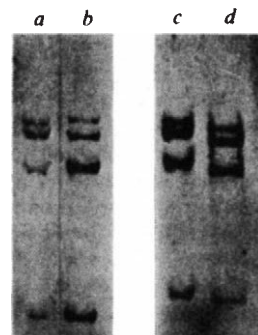


Fig. 3 Analysis of encapsidated RNA. Infected protoplasts ($\sim 1 \times 10^6$) collected 24 h after inoculation were homogenized in 500 mM sodium acetate, 80 mM magnesium acetate pH 4.5, using a ground-glass tissue homogenizer. Unlabelled BMV (200 µg) was added as a carrier and the sample was subjected to differential centrifugation to isolate virions¹⁴. The virus pellet was resuspended in 400 µl of a 1:10 dilution of the above buffer. RNA extraction and electrophoresis were as described in Fig. 1 legend except that 1M sodium salicylate was used as the autoradiographic enhancer for lanes c and d. Mock-inoculated protoplasts contained no radioactivity. a, c, d, RNA isolated from virus extracted from BMV-infected protoplasts cultured in ³H-Tyr; b, BMV Tyr-RNA charged *in vitro*. RNA contained in lane a was from a separate experiment than that contained in c and d. RNA in d was heated in TBE buffer (as in Fig. 2).

detectable levels of radioactivity. Figure 3 shows that all four encapsidated RNAs were radioactive—this was surprising as comparisons of the acceptor capacity for batches of RNA isolated in acidic or alkaline conditions from virions suggested that encapsidated BMV RNA was not tyrosylated². Therefore, the RNA was tested for the ability to lose radioactivity by deacylation in alkaline conditions. Figure 3c,d shows that most of the radioactivity remained after heating, suggesting that the radioactivity associated with the encapsidated RNA was not due to an aminoacylated tyrosine. It is likely that turnover of the ³H-Tyr during the 24 h incubation period resulted in the formation of ³H-labelled nucleotides that were subsequently incorporated into viral RNA, and thus the radioactivity was relatively stable to hydrolysis.

These results demonstrate that non-encapsidated viral RNA is aminoacylated within a natural host cell. The tyrosylated viral RNA appeared to be full-length as it co-migrated in polyacrylamide gels with full-length BMV RNA. This differs from the situation described by Joshi *et al.*¹⁶ for aminoacylation of turnip yellow mosaic virus RNA with valine in *Xenopus laevis* oocytes, in which the aminoacylated RNA appeared to have been cleaved into a fragment similar in size to tRNA; no full-length charged viral RNA could be detected. This discrepancy may reflect differences between plant and animal cells or possibly differences in methods of detecting charged RNAs.

Subgenomic RNA4 of BMV and two putative subgenomic messengers of BSMV were found to be aminoacylated. If aminoacylation is involved in RNA replication, such as the mechanism postulated by Hall and Wepprich¹² wherein aminoacylated viral RNA binds elongation factor and subsequently replicase, these results suggest replication of the subgenomic RNAs via their own templates.

We thank Dr A. O. Jackson for 'black hullless' barley seed and BSMV Type culture and for useful discussions. We also thank P. A. Kiberstis for the gift of *in vitro* aminoacylated BMV Tyr-RNA. This work was supported by NIH grant AI 11572 and NSF grant PCM 78-22526.

Received 26 January; accepted 8 June 1982.

1. Ahlquist, P., Dasgupta, R. & Kaesberg, P. *Cell* **23**, 183-189 (1981).
2. Hall, T. C. *Int. Rev. Cytol.* **60**, 1-26 (1979).
3. Dasgupta, R. & Kaesberg, P. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4900-4904 (1977).
4. Kohl, R. & Hall, T. C. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2682-2686 (1977).
5. Hall, T. C., Shih, D. S. & Kaesberg, P. *Biochem. J.* **129**, 969-976 (1972).
6. Agranovsky, A. A., Dolja, V. V., Gorbulev, V. G., Kozlov, Yu. V. & Atabekov, J. G. *Virology* **113**, 174-187 (1981).
7. Lane, L. C. & Kaesberg, P. *Nature new Biol.* **232**, 40-43 (1971).

8. Jackson, A. O. & Brakke, M. K. *Virology* **55**, 483–494 (1973).
9. Lane, L. C. *Virology* **58**, 323–333 (1974).
10. Shih, D. S., Kaesberg, P. & Hall, T. C. *Nature* **249**, 353–355 (1974).
11. Chen, J. & Hall, T. C. *Biochemistry* **12**, 4570–4574 (1973).
12. Hall, T. C. & Wepprich, R. K. *Ann. Microbiol.* **127A**, 143–152 (1976).
13. Dasgupta, R., Ahlquist, P. & Kaesberg, P. *Virology* **104**, 339–346 (1980).
14. Loesch-Fries, L. S. & Hall, T. C. *J. gen. Virol.* **47**, 323–332 (1980).
15. Gustafson, G. D. *et al. Virology* (in the press).
16. Joshi, S., Haenni, A. L., Hubert, E., Huez, G. & Marbaix, G. *Nature* **275**, 339–341 (1978).
17. Shih, D. S., Lane, L. C. & Kaesberg, P. *J. molec. Biol.* **64**, 353–362 (1972).
18. Bockstahler, L. E. & Kaesberg, P. *J. molec. Biol.* **13**, 127–137 (1965).

Differential scattering of circularly polarized light by the helical sperm head from the octopus *Eledone cirrhosa*

Marcos F. Maestre*, Carlos Bustamante†, Thomas L. Hayes*, Juan A. Subirana‡ & Ignacio Tinoco Jr†

* Donner Laboratory, Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720, USA

† Chemistry Department, Laboratory of Chemical Biodynamics, University of California, Berkeley, California 94720, USA

‡ Departamento de Química Macromolecular, Escuela Técnica Superior de Ingenieros Industriales, Barcelona (28), Spain

Helices are expected to scatter circularly polarized light preferentially. According to theory, the scattering depends on the pitch, radius and sense of helix^{1–4}; however, experimental data have been lacking. We now report quantitative measurements of the differential scattering of left and right circularly polarized light from the sperm head of the octopus *Eledone cirrhosa*. This sperm head is a left-handed helix with pitch and diameter of about 1 μm . It preferentially scatters left circularly polarized light versus right circularly polarized light in the forward direction, with a maximum differential scattering of 2% at 35° to the incident beam (wavelength = 0.442 μm). The preferential forward scattering of left circularly polarized light is consistent with theory^{1–4}. As a differential scattering of 0.1% is easily measured, this can be a useful new method to characterize helical biological objects.

Circular dichroism (CD) studies of complex biological macrostructures sometimes produce anomalies in the measured spectra. CD signals are observed outside the absorption bands of the chromophores, thus making the interpretation of the spectra difficult^{5–7}. These artefacts have been shown to arise from a difference in scattered intensities when left and right circularly polarized light is incident on the sample^{8–11}. This differential scattering produces a differential extinction, which is recorded by the CD spectrophotometer as if it were differential absorption. Experimental techniques exist to characterize these spurious scattering effects and to eliminate them^{9,12}. However, recent theoretical work^{1–4} has shown that the differential scattering contains detailed structural information. The theory shows that the angular dependence of the circular

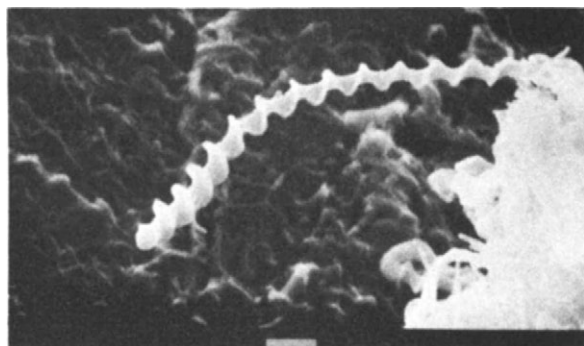


Fig. 1 Scanning electron micrograph of a freeze-dried sperm head from the octopus *Eledone cirrhosa*. The sperm head is a left-handed helix with a pitch of about 0.65 μm . Scale bar, 1 μm .

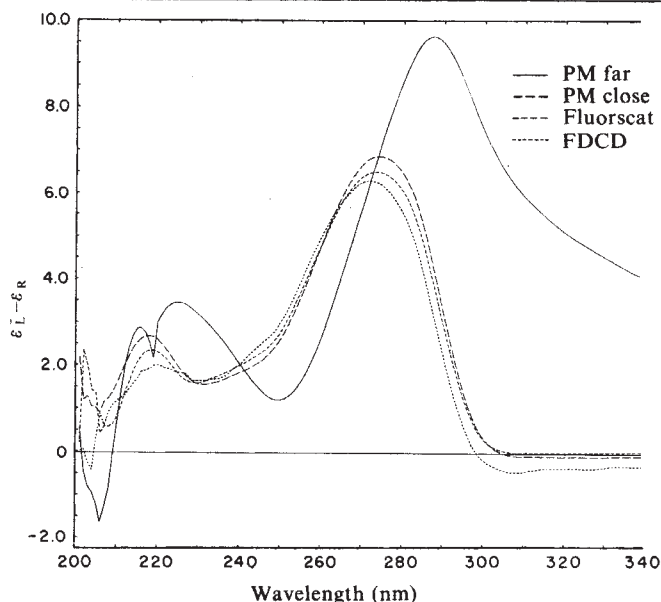


Fig. 2 Circular dichroism spectra in arbitrary units of *E. cirrhosa* sperm heads, measured by techniques having different solid angles of detection of scattered light. —, 'PM far' is the conventional configuration; — —, 'PM close' has the photomultiplier detector close to the sample to detect most of the forward scattered light; ····, fluorscat method; ····, fluorescence detected circular dichroism, FDGD, method. Sample contained 1 mg lyophilized heads in 10 ml of a 0.1 M NaCl, 10 mM Tris, pH 7.5 buffer.

intensity differential scattering, $\text{CIDS} = (I_L - I_R)/(I_L + I_R)$, reflects the structure of the chiral macromolecule. Here I_L and I_R are the intensities of the scattered light at a given angle when left and right circularly polarized light is incident. Application of the general theory to helices shows that polar plots of CIDS versus scattering angle give lobes whose sign, number and position are characteristic of the sense, pitch and radius of the helix.

We have built an instrument to measure the angular dependence of differential scattering of circularly polarized light (CIDS), and report some preliminary results that indicate the usefulness of such measurements. The nuclei of sperm cells from the octopus *E. cirrhosa* were chosen for initial study, because their large helical structure ensures large CIDS values, and their chirality and dimensions could be measured independently. Figure 1 illustrates the remarkable size and complexity of the helical head of the mature sperm from *E. cirrhosa*. These rigid nuclei have been measured by light microscopy to be 43 μm long, and to form a left-handed helix with a pitch of 1 μm and radius of 0.5 μm (ref. 13). Our measurements in the scanning electron microscope gave values for the pitch of 0.65 μm , with an outside diameter of 0.60–0.65 μm and an inner diameter of 0.25 μm . The difference in the two sets of values probably reflects the difference in techniques. On suspending purified nuclei in a buffer solution, intense light scattering is observed. Figure 2 shows the CD spectrum of the suspended nuclei. The curve labelled 'PM far' is the usual experimental arrangement, in which the photomultiplier mainly detects the transmitted beam; only light scattered very close to this beam is also detected. Note that the CD signal is very strong at long wavelengths, where the components of the nuclei (DNA and protein) do not absorb light, suggesting a large contribution from differential scattering of circularly polarized light. Of particular interest is the fact that this differential scattering effect is essentially eliminated by bringing the photomultiplier close to the sample and thus increasing the angle of acceptance of the scattered light (Fig. 2, curve labelled 'PM close').

Further small corrections were achieved by use of fluorscat cuvettes⁸ and fluorescent detected circular dichroism (FDGD) techniques⁹. Fluorscat is a double-chambered cuvette with the outer chamber filled with a fluorescent solution that captures all light not absorbed by the sample, except for a small sector

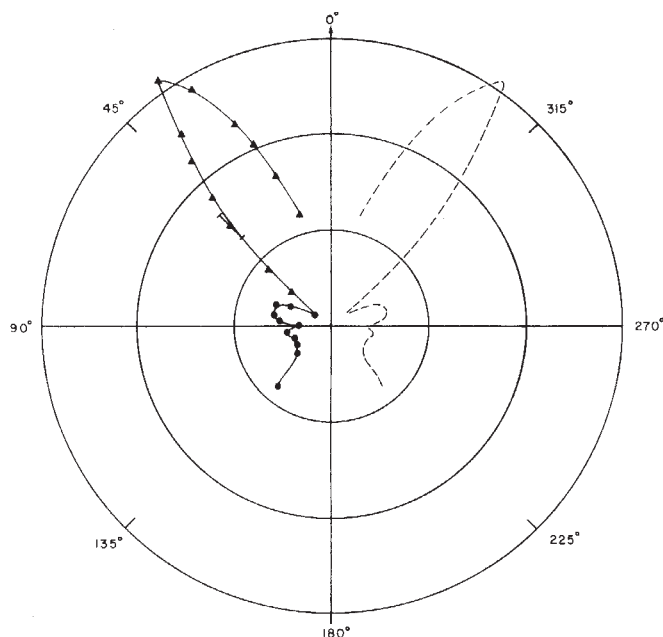


Fig. 3 Plot of CIDS $[(I_L - I_R)/(I_L + I_R)]$ versus scattering angle for *E. cirrhosa* sperm heads at a concentration of 0.25 mg ml^{-1} in the buffer of Fig. 2; the arrow illustrates the direction of incident beam (wavelength = 442 nm). $\blacktriangle$, Positive values; $\bullet$, negative values; —, measured data; ---, constructed by symmetry. The plot has a maximum positive value of 1.4×10^{-2} at about 35° ; the positive value means left circularly polarized light is preferentially scattered.

in the back direction of the scattering envelope. The FDCD technique involves the use of a fluorescer molecule that is not optically active and that does not bind to the particle being measured. Each scattering particle is now surrounded by a cloud of fluorescent molecules which measure all the light that is not absorbed, including light scattered backwards along the incoming light beam. Considering the minor corrections to the measured CD by the fluorscat and FDCD techniques (Fig. 2), we conclude that the sperm heads mainly scatter light differentially in the forward direction.

The apparatus used to measure CIDS directs linearly polarized light from a He-Cd laser (442 nm) through a Pockels' cell (an electro-optical quarter wave plate) with an applied square voltage pulse which produces alternately pure left and right circularly polarized light. The circularly polarized light is incident on the sample cuvette, and the scattered radiation is measured by a detector mounted on a goniometer. The signal from the detector is processed by a modified Cary 60 spectrophotometer and analysed in terms of CIDS.

Figure 3 shows the resulting polar plots of CIDS versus scattering angle for the sperm heads; a reference sample of polystyrene beads shows no differential scattering in the same conditions. This plot shows that the CIDS does occur mostly in the forward scattering direction (with a maximum at 35° from the incident beam), and also that left circularly polarized light is scattered more intensely than right circularly polarized light, that is, the forward lobe is positive. There is evidence of small negative differential scattering to the sides (90°) as is predicted from the minor correction of the fluorscat method (Fig. 2). The results presented here are preliminary and have been obtained with a maximum angular resolution of 5° ; however they are very reproducible. We are now working to improve the resolution of the apparatus to obtain more detailed information about the angular dependence of the differential scattering intensities.

The data shown in Fig. 3 are of the type expected from theory for a helix whose pitch or radius is larger than, but similar to the wavelength of incident light. The preferential forward scattering of left circularly polarized light indicates a left-handed helix. We have measured differential scattering from T4 and

T7 bacteriophage¹⁴ which are very different from each other and from the octopus sperm. We think that differential scattering of circularly polarized light can be very useful in determining the packing of DNA in bacteriophages^{15,16}, and the higher order structures of DNA in chromosomes¹⁷.

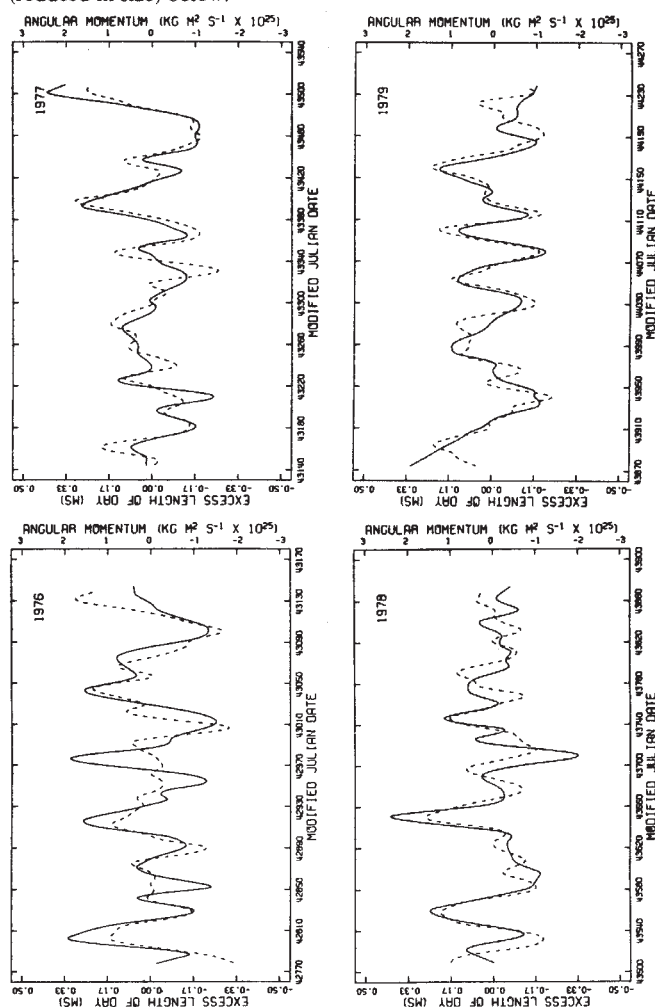
We thank Dr Phillip Burns for help with some of these measurements. This work was supported by NIH grants AI08247(M.F.M.) and GM 10840(I.T.) and by the Office of Health and Environmental Research of the US Department of Energy under contract 98, DE-AC03-76SF00098.

Received 22 April; accepted 11 June 1982.

1. Bustamante, C., Maestre, M. F. & Tinoco, I. Jr *J. chem. Phys.* **73**, 4273–4281 (1980).
2. Bustamante, C., Maestre, M. F. & Tinoco, I. Jr *J. chem. Phys.* **73**, 6046–6055 (1980).
3. Bustamante, C., Tinoco, I. Jr & Maestre, M. F. *J. chem. Phys.* **74**, 4839–4850 (1981).
4. Bustamante, C., Tinoco, I. Jr & Maestre, M. F. *J. chem. Phys.* **76**, 3440–3446 (1982).
5. Urry, D. W., Massoti, L. & Krivacic, J. *Biochem. biophys. Res. Commun.* **41**, 521–524 (1970).
6. Schneider, A. S., Schneider, M. J. & Rosenheck, K. *Proc. natn. Acad. Sci. U.S.A.* **66**, 793–798 (1970).
7. Maestre, M. F., Gray, D. M. & Cook, R. B. *Biopolymers* **10**, 2537–2553 (1971).
8. Dorman, B. P. & Maestre, M. F. *Proc. natn. Acad. Sci. U.S.A.* **70**, 255–259 (1973).
9. Reich, C., Maestre, M. F., Edmonson, S. & Gray, D. M. *Biochemistry* **19**, 5208–5213 (1980).
10. Tinoco, Jr., I., Bustamante, C. & Maestre, M. F. *A. Rev. Biophys. Bioengng* **9**, 107–141 (1980).
11. Maestre, M. F. & Reich, C. *Biochemistry* **19**, 5214–5223 (1980).
12. Dorman, B. P., Hearst, J. E. & Maestre, M. F. *Meth. Enzym.* **27**, 767–796 (1973).
13. Maxwell, W. L. *Proc. R. Soc. B186*, 181–190 (1974).
14. Tinoco, I. Jr, Bustamante, C. & Maestre, M. F. in *Structural Molecular Biology: Methods and Applications* (eds Davies, D. B., Saenger, W. & Danyluk, S. S.) (Plenum, New York, 1982).
15. Tikhonenko, T. I. *Comprehensive Virol.* **5**, 1–117 (1975).
16. Black, L. W. & Silverman, D. J. *J. Virol.* **28**, 643–655 (1978).
17. Sedat, J. & Manuelidis, L. *Cold Spring Harb. Symp. quant. Biol.* **42**, 331–350 (1977).

Erratum

In the letter 'Atmospheric angular momentum and the length of day: a common fluctuation with a period near 50 days' by R. B. Langley *et al.* *Nature* **294**, 730–732 (1982), the curves in a panel of Fig. 1 were rotated through 180° . The figure is shown correctly in reprints and (reduced in size) below.



BOOK REVIEWS

Blunders in the business of risk

Dorothy Nelkin

In 1885 a British tourist in the United States was travelling on a train when it began to shake violently. Panicked, he asked the conductor to stop the train, but to no avail. The train jumped its tracks and most of the cars were smashed. Yet to his amazement, his fellow passengers applauded the conductor's efforts to arrive on time. So consumed were they by the idea of progress, they failed to recognize a near fatal disaster.

Clearly, attitudes have changed. Today, the United States has been described as a "risk-averse" society, and some feel that our preoccupation with risks is threatening our commitment to technological progress and economic growth. In this environment the field of risk assessment has become a veritable industry. The assessment, management and adjustment to risk is estimated to cost between \$200 and \$300 billion per year, that is, from 10 to 15 per cent of the gross national product. The fear of risk is reflected in legislation and regulation. A programme within the National Science Foundation supports research on risk, hoping to provide a rational and reasonable basis for technical decisions. The availability of funds has encouraged a proliferation of consulting firms and academic research programmes. Indeed, the field of risk assessment, engaging geographers and engineers, economists and systems analysts, psychologists and anthropologists, has replaced technology assessment as the fashion of the day.

Risk analysts are preoccupied with two central questions: "How safe is safe enough?" and "What factors influence perception of risk?". Systematic efforts to estimate and evaluate the risks inherent in different technologies and then to quantify the acceptability of risk began to burgeon in response to environmental controversies, in particular over nuclear power. Hoping to develop a "rational" means to make decisions about hazardous technologies, a number of scholars from the field of engineering safety analysis sought to establish a quantitative basis to compare risks from different technological choices and to calculate their relative costs and benefits. They were encouraged by legislation requiring the risks of technological choices to be systematically weighed against social benefits. Their analytical approaches to risk, based on engineering and economic models, included a variety of forecasting and systems analysis techniques. They confronted

Risk and Culture: An Essay on the Selection of Technological and Environmental Dangers. By Mary Douglas and Aaron Wildavsky. Pp.224. ISBN 0-520-04491-6. (University of California Press: 1982.) \$14.95, £11.25.

serious methodological dilemmas. Indeed, the uncertainty about the nature and extent of the risks inherent in many technological choices seems to defy systematic and objective analysis. In any case, the various efforts to calculate the risks and benefits of technology proved to be poor indicators of the political and social acceptability of risk.

Difficulties in predicting public attitudes focused attention on the psychological and attitudinal factors that enter subjective perceptions of risk. Psychologists began to study risk perception through laboratory experiments, gaming situations and survey techniques in order to identify the subjective values that guide individual choices. Their work on the characteristics of risk that influence judgements about acceptability suggests that risks that are involuntary, uncertain, unfamiliar and potentially catastrophic are most difficult for people to accept. In their studies, psychologists try to arrive at a quantitative determination of risk acceptance. However, limited by their focus on the individual, they often fail to account for variations in the perception of risk in different social and cultural contexts. Why indeed did the British tourist perceive risks so differently from his fellow travellers?

These social dimensions of risk perception are the focus of *Risk and Culture*. Seeking to explain the preoccupation with risk in American society, the authors start with a provocative and promising premise — that it is inappropriate to define simply the problem of risk either in terms of objectively calculated physical risks or subjectively biased individual perceptions. Between subjective perceptions and physical science, they argue, lies a culture; and perceptions of risk are necessarily embodied in the complex system of beliefs, values and ideals which constitute this culture. This premise carries the debate about risk beyond the narrow technical questions of safety and into the domain of values and social choices.

This view of risk as a socially constructed phenomenon is a creative and refreshing addition to the risk analysis literature. It could lead to research into a variety of questions, shedding light on the social and political dimensions of risk perception.

However, instead, the authors carry their argument in a direction that reflects their personal concerns with challenges to the *status quo*.

Their argument is as follows. Risk is a kind of euphemism for social and moral issues. Both modern and primitive societies select dangers to buttress their view of the social system. Risk becomes an explanation of failure, a way of attributing blame, a moral judgement about conduct inspired by a vision of the social and moral order. People predisposed to prefer different forms of social organization will perceive different dangers. Douglas and Wildavsky thus shift the critical question in risk analysis from "What are the real risks?" to "Why do different people perceive risks in different ways?" — "What issues of accountability or values are at stake?"

They build their argument by rejecting the conventional assumptions of most risk analyses. They dispose of explanations based on scientific evidence, pointing to the disputes among scientists and the real uncertainties surrounding many questions of risk. They observe how scientific judgements themselves are rooted in social expectations. Similarly, they account for the rigid categories of voluntary and involuntary risk underlying "objective" analyses by observing how the choices of risk assessment methods are themselves biased by underlying cultural assumptions. Then they turn to their own analysis, suggesting how different cultures with shared values and social institutions will select or highlight certain risks and minimize others.

The argument is built on Douglas's superb work on "pollution" — the idea that "pollution beliefs" are used by all societies to define boundaries that will preserve social categories and political relationships. It is a convincing thesis. The very language of those concerned with risks vividly reveals their social and political orientation. Why else would critics of nuclear technology talk of "electro-fascism" or "nucleocracy"? However, my enthusiasm for this approach to risk begins to wane when Douglas and Wildavsky develop a three-part typology of "risk portfolios" in American society. They assert that the evaluation of risk is based on (i) hierarchical assumptions, (ii) market or individualistic rationality, or (iii) sectarian values. The first two comprise a "centre" in which maintenance of the existing system is the primary goal. "Sectarian" rationality comprises a border, remote from the centre of power. The perspectives

of sectarian groups on risks reflect their rebellion against the centre and their own needs to maintain group allegiance. Their views are marked by "disdain of the world, fear of pollution, of cancerous contamination" and are propelled by "moral fervor". The sectarian "wins adherence if he can threaten bigger dangers and associate them with the corruption of the system".

The authors develop their model of the sectarian society less from evidence than from their own opinion of the radical wing of the environmental movement. They quickly lapse into pejorative language that hardly masks their disdain. They portray such groups as self-indulgent and living in permanent and nagging opposition to the centre with no intention of assuming responsibility or power. And they denigrate the strengths of sectarian ideals suggesting that, when not allowed to enjoy government subsidy, such groups simply become part of the very system they reject.

The authors feel that such sectarian values are increasingly important in American society and have won "extraordinary success". They do, indeed, account for our preoccupation with risk. There are, however, many problems with their argument. After all, the "sustained assault" by public interest and environmental groups lasted but a decade. In the context of today's politics the "extraordinary success" of environmental groups as compared to business interests is hardly convincing. Moral fervour in the United States today comes from other directions, indeed from the centre itself. Concerned about challenges to the hierarchical and market values that constitute the centre, the authors ignore the usefulness of sectarian groups as a source of constructive criticism and countervailing power. They neglect, for example, the positive contributions of "border groups" such as the alternative technology movement who have put forward specific proposals for minimizing risk. And to suggest that such sectarian — i.e. powerless — groups could gain power if they really wanted it, is to totally ignore political and economic constraints. In their criticism of such groups, the authors describe "lobbying, litigation and non-violent obstruction" as "border strategies", forgetting that these are also conventional political actions exploited by bureaucracies and businesses.

The authors also dismiss the possibility that the reality of risk has changed in any significant way, essentially ignoring the increased scale and changing character of technology-induced risk. Associating the fear of risk with radical groups, their theory has no place for those who work in genuinely dangerous jobs, or the residents of Love Canal, or the 700,000 middle-class people who converged in New York City to protest against the risk of nuclear war.

A cultural theory of risk could be developed in a different direction with

greater emphasis on the social position and the special vulnerability of certain groups with limited economic or political choice. It could call attention to the experience and perspectives of people who deal with risks as a fact of everyday life. Clearly judgements of risk reflect cultural values. But after all there is a social reality. If indeed the issue of risk has become an arena for disputes over power, why not take off from this approach to examine critically corporate perceptions and behaviour with respect to risk? Why not ask how factors at work and in daily life influence people's risk perceptions?

Douglas and Wildavsky end their book by frankly stating their own bias towards the centre. Basing their analysis on the acceptance of given institutional arrangements, they, along with other risk analysts, have masked many problems by avoiding critical questions about the very political and social relationships that they themselves admit are fundamental to the selection and perception of risks. □

Dorothy Nelkin is a Professor in the Cornell University Program on Science, Technology and Society and the Department of Sociology.

An enlarged Universe

Owen Gingerich

The Expanding Universe: Astronomy's 'Great Debate' 1900–1931. By Robert W. Smith. Pp.220. ISBN 0-521-23212-0. (Cambridge University Press: 1982.) £19, \$29.50.

ONE of the extraordinary developments of twentieth-century science has been the recognition and measurement of the great size of the Universe. The most dramatic moments in this revolutionary understanding took place from 1918 to 1924, primarily at the great mountain observatories in the American West, and a focal event was the debate on the scale of the Universe between Harlow Shapley and H.D. Curtis in Washington, in June 1920.

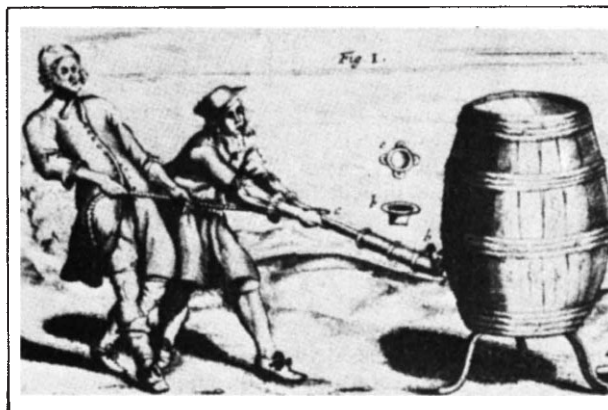
The Shapley–Curtis debate has achieved

something of near mythological status among astronomers. Robert Smith capitalizes on the mystique of the specific debate in his subtitle to extend the argumentation over a broader time span; and well he might, for, as he informs us, the actual encounter before the National Academy of Sciences left a great deal to be desired as far as scientific depth was concerned. The main title is likewise stretched beyond its specific technical connotation, and in an almost punning fashion refers more to a Universe ever larger in man's conception than in physical expansion.

The Shapley–Curtis–Hubble era has attracted a variety of historical treatments in the past decade. Smith's in-depth research fills in the contemporary details too often blurred by our 20–20 hindsight. For example, the status of the globular clusters as possible external galaxies has been almost entirely neglected in modern historical accounts of the 1915–1925 period. Similarly, he fills in the story of the first faltering attempts in the 1920s to derive a relation between the velocities and distances of the spirals, without which certain aspects of Edwin Hubble's approach to the problem can scarcely be comprehended.

In giving us a near-definitive account of the fascinating and turbulent growth in astronomical knowledge, Smith has assiduously avoided the historian's sin of evaluating the past with today's standards. Yet it would have been illuminating, I think, for him to have allowed that Shapley's joining of the period–luminosity relations for cluster-type variables and the Cepheids, for which he was criticized at the time, was in fact incorrect. Equally, Shapley's own criticism of Hubble's assumptions about the similarity of certain galaxies and their components has today proved valid. In the event both Shapley and Hubble pushed on with their faulty programmes, to the considerable advance of science in their day. This may well be telling us something interesting about the nature of scientific progress, but Smith's strictly historical approach prevents any such speculation. □

Owen Gingerich is a Professor of Astronomy and History of Science at the Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts.



An experiment of Otto von Guericke (1602–1686), a predecessor of Robert Boyle, attempting to produce a vacuum by exhausting a sealed cask filled with water. The illustration is taken from a corrected reprint of A. Rupert Hall's *From Galileo to Newton*, first published in 1963. The new edition is published by Dover in the United States and Constable in Britain, prices \$6, £4.50 respectively (pbk).

Polymeric molecules in motion

Warner L. Peticolas

The Theory of Vibrational Spectroscopy and Its Application to Polymeric Materials. By Paul C. Painter, Michael M. Coleman and Jack L. Koenig. Pp.530. ISBN 0-471-09346-7. (Wiley: 1982.) £44.50, \$80.

A DISTINGUISHED figure in the field of vibrational spectroscopy, Professor Jack Koenig of Case Western Reserve University has collaborated with two of his younger colleagues to produce a book which should be, for some time to come, the definitive work on the application of the theory of molecular vibrations to polymeric materials. Starting with a clear treatment of the classical mechanics of the vibrations of small molecules using both Cartesian and internal coordinates, the book takes the reader in a logical way to calculations of the normal modes of vibrations of polymer chains.

Due to the exigencies of space, some short cuts have been taken. For example, although the book starts with a discussion of small molecules, there is no mention of the formulae for calculating molecular internal coordinates from molecular Cartesian coordinates. Instead, the general concept of internal coordinates and their relationship to Cartesian coordinates is worked out in detail using matrix techniques. If the reader would like to work out the G matrix for small molecules with a desk calculator, he will have to turn to one of the listed references.

However, one gets the feeling that this book is really for readers who wish to do their calculations using computers. With a good computer program, it is possible to carry out normal mode calculations without actually going into a detailed derivation of the precise mathematical formulae for calculating each type of internal coordinate. Computer programs are available which define the internal coordinates and then immediately give instructions for calculating them from the coordinates of the atoms, which must be supplied as data. This approach is stressed in Chapter 7, entitled "Solving the Vibrational Problem Using Computer Methods". The authors note that it is difficult to use a desk calculator to solve secular matrices whose dimensions are greater than 3 by 3 and that it is much easier to go directly to a computer. They conclude their discussion of simple molecules with a detailed analysis of the benzene molecule. This is an area to which they themselves have made an important contribution in elucidating the vibrational spectra of benzene and its two important carbon-13 isotopic derivatives, as well as those of numerous deuterium substitutive derivatives.

After giving a brief account of the quantum mechanical description of infra-

red and Raman intensities, the authors discuss the lattice dynamics of chain molecules. This is the strongest feature of the book. Starting with the simple linear diatomic chain, and going to a detailed mathematical discussion of the Wilson-GF method, they present a clear method for calculating the normal modes of helical polymeric chains. The theory is followed by consideration of the normal modes calculations of many common polymeric

materials. The book closes with an introduction to biological polymers which includes a discussion of the normal modes of polyglycine I.

Anyone wishing to do research in the field of vibrational spectroscopy of polymeric materials, including biological polymers, or anyone interested in the force fields of macromolecules, will find this book a time-saving, helpful introduction to a very complex but fascinating subject. □

Warner L. Peticolas is a Professor of Biophysical Chemistry at the University of Oregon, Eugene.

Alternative hosts for genetic engineers

Noreen Murray

Gene Cloning in Organisms Other than E. coli. Vol.96 *Current Topics in Microbiology and Immunology.* Edited by P. H. Hofschneider and W. Goebel. Pp.259. ISBN 0-387-11117-4/3-540-11117-4. (Springer-Verlag: 1982.) DM92, \$42.90.

BIOCHEMICAL approaches to the molecular cloning of DNA became simple when, in 1972, it was shown that DNA digested with certain endonucleases yielded discrete fragments with cohesive ends. In the intervening decade, molecular cloning techniques using *E. coli* K12 as host have led to a detailed understanding of the mechanism of action of several prokaryotic genes and have revolutionized our concept of the organization of the eukaryotic genome. The full impact of this technology cannot, however, be realized just by cloning, sequencing and changing genes. A complete understanding of gene expression requires the cloned genes to be studied within the environment of their provenance, whilst commercial production of polypeptides may demand hosts free of endotoxins from which products can be exported across the cell membrane.

Volume 96 of *Current Topics in Microbiology and Immunology* surveys the current and potential use of *Bacillus subtilis*, pseudomonads, streptomycetes, fungi, animal and plant cells as hosts for molecular cloning. In many instances a more complete view of any one genetic system is achieved by including two chapters covering one system. For each system the following topics are documented: vectors; uptake of DNA into cells; the recognition of cells carrying recombinant DNA molecules; and the fate of the donor DNA following its uptake.

The development of alternative prokaryotic systems inevitably draws heavily upon experience with *E. coli* where the concepts of vector construction are well established. Consequently, novel information is often relatively specialized and likely to be of most interest to those readers

who wish to use bacteria other than *E. coli* as hosts for cloning experiments. For *Streptomyces* much useful technical information, including laboratory methods and recipes, is provided. Of more general concern are the justifications for studying different bacterial species and the accounts of experiments relevant to the barriers preventing expression of heterologous genes.

The fungi are represented by a brief article documenting the beginnings of molecular cloning in *Neurospora crassa* and two chapters conveying some of the dramatic advances achieved in the molecular biology of *Saccharomyces*. The second of these two articles is in no way redundant, but the first (by Hinnen and Meyhack) is so clearly presented that it should be enjoyed by students of both genetics and molecular biology.

I approached the contributions on higher cells as an interested outsider, anxious to be informed of some of the more exciting developments in molecular biology. In the event I found them somewhat disappointing. The chapter on selectable markers for gene transfer was indeed useful, covering the systems and their use in transformation with some critical evaluation of the resulting studies of gene expression. However, a detailed diagram summarizing the relevant features of the Herpes TK gene and a chart indicating those aspects of nucleotide metabolism of relevance to the selective systems in use would have been welcome. In contrast, the short chapter on animal viral vectors, principally SV40, presents only a limited account of the field. Perhaps the authors anticipated the inclusion of a second and complementary article; certainly the space devoted to animal viral vectors is only a half or a third that for either *Bacillus subtilis* or pseudomonads. Unfortunately the volume was conceived too early to take account of recent success in the transformation of *Drosophila* cells.

A detailed survey of liposomes as a

NEW

FORAMINIFERA

J. R. HAYNES

Reader in Geology
University College of WalesDecember 1981; £50.00;
480pp; 34pp plates; ISBN 0
333 28681 2

CONTENTS

Preface; The Scope of Foraminiferal Studies; Collection, Preparation and Examination (including Annotated Literature on Technical Methods); The Living Foraminifer; Test Morphology and Composition; Classification of the Foraminifera; The Agglutinating Foraminifera; The Fusulinida – The Miliolida; The Nodosariida; The Buliminida; The Robertinida; The Rotaliida (Smaller); The Rotaliida (Larger); The Globigerinida; References; Index.

This work is an advanced text covering the biology, morphology and classification of the Foraminifera and their use in stratigraphy and palaeoecology.

Superbly illustrated, the text draws on the author's experience in oil exploration and twenty years' teaching at both undergraduate and postgraduate level.

The main features of this work are:

- A new classification, illustrated by hand-drawn figures of the main genera, selected scanning electron micrographs and photomicrographs of thin sections.
- New theories linking the morphology of the foraminiferal shell with its adaptive function.
- The integration of the study of each of the orders which tend to succeed each other in time to illustrate their use in biostratigraphy and palaeoecology from the Cambrian to the Recent.
- It brings the student up against current problems and the dangers of an over-simplified approach to palaeoecology, evolutionary studies and biostratigraphy, and synthesises the result of recent research which is at present widely scattered in the literature.

For further information write to
FRANCES ROACH, Product Manager,
Scientific & Medical Division,
Macmillan Press Ltd., Houndmills,
Basingstoke, Hampshire RG21 2XS,
U.K.

SCIENTIFIC
& MEDICAL

M

MACMILLAN
PRESS LTD

carrier system is included but there is no chapter devoted either to microinjection, or protoplast fusion, as a means of DNA or RNA transfer. Plant systems are represented by comprehensive accounts of cauliflower mosaic virus and Ti plasmids of *Agrobacterium*. The former includes as appendices a computer print-out of the viral DNA sequence and derivative information.

The broad topic of genetic engineering in hosts other than *E. coli* should appeal to a wide audience, and I would hope to find a copy of this useful reference material in all university libraries. □

Noreen Murray, currently on leave of absence in the European Molecular Biology Laboratory, Heidelberg, is Reader in the Department of Molecular Biology at the University of Edinburgh.

Pollution in context

Eirene White

Principles of Pollution Control. By Francis Sandbach. Pp.174. Pbk ISBN 0-582-30042-8. (Longman: 1982.) £5.95, \$11.95.

THIS is not a book addressed to the serious natural scientist. To be fair, it does not pretend to be. It is the first study in a series entitled "Themes in Resource Management".

The manifold objectives of the series, as stated by the editor-in-chief, Professor Bruce Mitchell, are, first

to identify and examine enduring resource management and development problems [and thence] to assess responses to these problems in a variety of world regions.

Of particular concern are research and action programmes. The former are to be evaluated in terms of the nature and adequacy of data, the theoretical and practical understanding displayed and the identification of special problem areas. Action programmes are to be reviewed in terms of their effectiveness in promoting "economic efficiency, social equity and environmental harmony".

This orientation [writes Professor Mitchell] should provide a departure point for considering the transferability of experience from one jurisdiction to another.

A final objective is

to explore the way in which resource analysis, management and development might be made more complementary to one another.

Set against these fairly demanding requirements, one may ask how far in his study of the principles of pollution control Francis Sandbach can be said to have complied with them, and in any case to what readership his book is directed. It would not satisfy a reader looking for a

scientifically-based account of pollution control over the past few decades or the scientific outlook for the future. But that is not what Mr Sandbach envisages. As a lecturer in interdisciplinary studies at the University of Kent, it is natural that he should seek to paint with a broad brush. One needs to go no further than the introduction to realize that although he has indeed read widely and assiduously about pollution in its various manifestations and provides a great deal of helpful factual information about how it is handled in Britain and elsewhere, he is really turned on by social and political theory rather than by science or technology.

Anyone who wants a schematic view of pollution control would do better to turn to Martin Holdgate's *A Perspective of Environment Pollution*, published by Cambridge University Press in 1979. Nevertheless, there is some validity in Francis Sandbach's thesis, as he forcefully illustrates in his case study of asbestos, the classic example of the clash of interests between health and profits. He devotes another chapter to a not very satisfactory attempt to discredit what he describes as the pluralist/behavioural perspective of those such as Eric Ashby and Mary Anderson who in their historical account of air pollution control in Britain (*The Politics of Clean Air*; for review see *Nature* 298, 207; 1982) pay less attention than he thinks fit to economic determinants.

The six preceding chapters comprise a rapid and necessarily superficial survey of risk/dose/response relationships; cost effectiveness; the alternatives of recycling or of using a less polluting technology; "the polluter pays" principle and the possible use of pollution taxes; the setting of standards or threshold limit values; the choice of strategies; and the administrative and legal framework for pollution control in the United Kingdom, the EEC and the United States, with glimpses of China, Japan and Scandinavia.

There is much of interest in the book, but it is all rather breathless, plucking a fact here and picking up a figure there. An extensive bibliography includes many useful references to periodical literature (including some that may be more emotive than scientific). The index, by contrast, is exceedingly thin and anyone wishing to look up Seveso or Minamata, for example, has to scan the text.

As a means of introducing students to the general problems of pollution control set in a comprehensive context, this study, in the hands of a well-informed lecturer, could be very helpful. For the more general reader, such as an elected member of a local authority, who might well be in need of guidance, it is likely to be too allusive and, at times, bewildering in its range of detail. □

Baroness White is a Member of the House of Lords Select Committee on Science and Technology (Water Sub-Committee), and was a Member of the Royal Commission on Environmental Pollution from 1974 to 1981.